PRAGMATIC FACTORS OF DEONTIC REASONING

STEPHEN GEORGE KILPATRICK BSc (Hons), MSc, CERT. ED.

A thesis submitted in partial fulfilment of the
requirements of the University of Wolverhampton
for the degree of Doctor of Philosophy

April 2007

Revised October 2008

Amendments & corrections March 2009

Declarations

Material included in Chapters 3 and 4 were presented at the following conferences:

Over, D.E., Manktelow, K.I. & Kilpatrick, S.G. (1998, June) Deontic reasoning; its philosophy & psychology. Paper presented at The Society for Philosophy and Psychology (SPP) Annual Meeting. University of Minnesota, Michigan, USA.

Manktelow, K.I., Kilpatrick, S.G. & Over, D.E. (1996, August) Pragmatic effects on deontic inference. Poster presentation at the Third International Conference on Thinking, University College London, London, UK.

Manktelow, K.I., Over, D.E. & Kilpatrick, S.G. (1996, July) Pragmatic factors in deontic reasoning. Paper presented at the XXVI International Congress of Psychology, Montreal, Canada.

Material in Chapters 4 and 5 was included in the following book chapter and conference:

Manktelow, K.I., Fairley, N., Kilpatrick, S.G. & Over, D.E. (2000). Pragmatics and strategies for practical reasoning. In W. Schaeken, G. De Vooght, A. Vandierendonck & G. D'Ydewalle (eds.), Deductive Reasoning and Strategies. Mahwah, NJ: Lawrence Erlbaum Associates.

Over, D.E., Kilpatrick, S.G. & Manktelow, K.I. (2001, June). Mitigation, aggravation, cheating and violations. Paper presented at the Human Behaviour & Evolution Society Conference, University College London, London, UK.

Material included in Chapter 6 was presented at the following conferences and forms the basis of a forthcoming journal article:

Kilpatrick, S.G., Manktelow, K.I. & Over, D.E. (2000, August) Power of source as a pragmatic factor of deontic reasoning. Poster presentation at the Fourth International Conference on Thinking, Durham University, Durham, U.K.

Kilpatrick, S.G., Manktelow, K.I. & Over, D.E. (2007) Power of source as a factor in deontic inference. *Thinking and Reasoning*, 13(3), 295-317.

Material included in Chapter 7 was presented at the following conferences:

Kilpatrick, S.G., Manktelow, K.I. & Over, D.E. (2003, July) Pragmatic effects and judgement revision in deontic reasoning. Paper presented at the Psychology Postgraduate Affairs Group (PSYPAG) Conference. University of Wolverhampton, UK.

Kilpatrick, S.G., Manktelow, K.I. & Over, D.E. (2003, September) Judgement revision with conditional promises, threats and warnings. Poster presentation at The XX BPS Annual Cognitive Psychology Section Conference. University of Reading, UK.

Acknowledgements

I wish to acknowledge the support and continuous encouragement of my two supervisors, Professor Ken Manktelow and Professor David Over. I had greater contact with Ken who provided much of the inspiration for the research reported in this thesis. I miss our lengthy chats in the Art Gallery and other places around Wolverhampton City Centre where ideas were discussed and developed. I must also acknowledge Ken's patience when progress was slow and endless deadlines were missed. Although I had less contact with David, his phenomenal output provided me with much theoretical material and aided me in keeping abreast of developments in the thinking and reasoning field. As will be seen David's work provided the theoretical basis to this thesis. On our infrequent meetings David was always cordial and treated me as an old friend. I can only describe the two as true Scholars and Gentlemen!

I am also indebted to my family who have had to endure this project over so many years. Carolyn has given me the space and time required to undertake such a mammoth task, even when she would prefer that I was doing something else. Our four children have also been tolerant of my absences and pre-occupations for which I am dearly grateful.

Dedication

This work is dedicated to my parents, particularly to the memory of my Mother, and to Carolyn and our four children.

Contents

<u>Abstract</u>

This thesis is concerned with pragmatic factors of deontic reasoning, namely scale of violation, aggravating and mitigating circumstances and power of source. Nine experiments are reported investigating deontic reasoning and judgement revision. Experiment 1 established scale of violation as a modifying factor of a *working rule* with an inferential reasoning task, however, the effects were not transferred to a deductive reasoning task in Experiment 2. Scale of violation and circumstances were found to influence the reasoning of motoring violations with a major offence and aggravating circumstances being rated as more serious and receiving greater fines than a minor offence or mitigating circumstances (Experiments 3 & 4). These effects were also observed with a judgement revision task (Experiment 5). Power of source was included as an additional pragmatic factor and was found to influence the reasoning of conditional statements (Experiment 6), inducements (Experiment 7) and ratings of credibility and probability of outcomes (Experiment 8). The final study (Experiment 9) found significant effects for scale of violation / compliance and power of source within a judgement revision task. However, no difference was observed in the reasoning of superordinate and non-superordinate statements. The findings are explained in terms of the conditional probability hypothesis.

Chapter 1

Deontic Reasoning

1.1 Introduction

Deontology has a long philosophical history but has only recently been studied by psychologists. The seminal paper of Cheng & Holyoak (1985) was the key influence with their insight that reasoning was most likely to be facilitated when a deductive reasoning task involved a permissible or obligational context, that is, a deontic context. This chapter will describe the precursory areas of research that lead to the psychological investigation of deontic reasoning. The effects of content and context in syllogistic reasoning and on the Wason selection task (Wason, 1968) will be reviewed before examining deontic reasoning in detail. The philosophical background to deontology will be explored, followed by an outline of the major psychological theories that attempt to explain this recent and developing branch of human reasoning.

1.2 Content & Context Effects in Syllogistic Reasoning

The traditional distinction in the study of reasoning has been between deductive and inductive reasoning. Deductive reasoning involves the derivation of conclusions from given premises; the premises must be assumed to be true and, therefore, the inferred conclusion must also be true. Inductive reasoning involves inferring conclusions based on ones knowledge or experience, thus, one goes beyond the given information, however, the drawn conclusion cannot be assured of being true. The focus in this section will be on deductive reasoning tasks, however, as Manktelow (1999) notes such a distinction may be erroneous as more recently the two areas of research have become blared.

9

For any deductive reasoning task a distinction can be made between its *form* and *content*; the form being the problems logical structure and the content being the scenario or context in which the problem is based (Evans, 1989). Evans (1989) distinguished between three general types of problem content; abstract, arbitrarily realistic, and knowledge-related content. Abstract problems are devoid of everyday knowledge or experience and such tasks are considered a test of pure reasoning. Problems comprising arbitrarily realistic material contain everyday terms and concepts but do not invoke relevant knowledge and beliefs to reach a solution and are, therefore, taken as problems requiring reasoning based on the logical form of the problem. Knowledge-related content problems utilise materials that link directly to an individual's previous knowledge, experience, and/or beliefs. Thus, it is not at all clear whether the form or content of the problem is influencing the process of reasoning. It is this latter type of problem content that will be the focus of this section and the next when we examine syllogistic reasoning and the selection task, respectively.

Aristotle considered *rationality* to be the ultimate human ability. Aristotle developed the syllogism and categorical logic and his notion of reason was based on the ability to think in a formal and systematic way. This logical explanation of reasoning remained unquestioned for several centuries. However, empirical investigations with reasoning tasks have shown that the content of materials and the context in which they are set influence obtained responses. Some of the research examining content and context effects in syllogistic reasoning will now be discussed along with their implications.

According to logic the *form* of an argument should determine ones reasoning, irrespective of *content*. Further, from a Piagetian perspective adults should have attained the formal operational stage of thinking, and any contextual influence would be indicative of the inferior concrete operational stage (Evans, 1982).

One of the earliest systematic investigations of syllogistic reasoning was conducted by Wilkins (1928). She used syllogisms comprising of a major and minor premise with participants choosing a conclusion from a given list. The syllogisms entailed both valid and invalid conclusions that were embedded in four types of content. The type of problem with the corresponding correct response is given below:

Problem Content	Correct Response (%)
Thematic	84
Abstract	76
Unfamiliar	75
Belief bias	80

Table 1.1 Percentage of correct responses to four types of syllogism (Wilkins, 1928; adapted from Evans, 1982)

Wilkins did not provide significance levels but Evans (1982) calculated that reasoning with thematic-rich syllogisms was significantly superior to reasoning with both abstract and unfamiliar materials and was marginally significant compared to syllogisms embedded in a belief bias context. Wilkins interpreted her findings in terms of belief bias. She found realistic (thematic) content could facilitate logical reasoning relative to abstract and unfamiliar content. Further, prior beliefs and attitudes could bias reasoning; hence the *belief bias effect*. The belief bias effect was supported by a number of contemporary researchers (Janis & Frick, 1943; Lefford, 1946; McGuire, 1960).

An alternative influential theory of syllogistic reasoning, at that time, was the *atmosphere hypothesis* (Woodworth & Sells, 1935). This hypothesis proposes that the premises create an *atmosphere* indicating a particular conclusion. For instance, a syllogism with a negative premise (No… or Some…Not) creates a negative atmosphere suggesting a negative conclusion, while a syllogism with an *extensional* premise (Some… or Some… not) creates an atmosphere in which an *extensional* conclusion is more likely to be accepted.

Research that has attempted to test the two proposals, the belief bias effect and the atmosphere hypothesis, has tended to support belief bias (Janis & Frick, 1943; Lefford, 1946; Morgan & Morton, 1944). However, much of this early empirical work has been criticised on methodological grounds (Evans, 1982; Revlin, Leirer, Yopp & Yopp, 1980).

Chapman & Chapman (1959) claimed that at least some errors in syllogistic reasoning could be explained by what they called the *conversion hypothesis*. Individuals are said to convert one or both premises of an argument and to reason from the converted premises. Statements in moods I (Some) and E (No) convert to logically valid inferences while statements in moods A (All) and O (Some…not) convert to logically invalid conclusions. The content of premises can influence conversion. For example, given the A statement, "All dogs are animals" individuals are unlikely to convert this to "All animals are dogs" because of prior knowledge concerning dogs and animals. However, the logically equivalent statement, "All ticket holders are allowed to enter the exhibition" is likely to be converted to, "All the people who are allowed to enter the exhibition are ticket holders". There exists a number of studies in support of the

conversion hypothesis, for instance Newstead (1990) found such errors common when participants were requested to say what follows from given statements.

Content effects have also been observed in areas other than syllogistic reasoning, including causal attribution tasks (Cheng & Novick, 1991), categorical syllogisms (Evans, 1982; Evans, Barston & Pollard, 1983) and conditional reasoning tasks (Cheng & Holyoak, 1985; Cheng, Holyoak, Nisbett & Oliver, 1986; Griggs & Cox, 1982; Manktelow & Evans, 1979; Newstead, Ellis, Evans & Dennis, 1997; Wason & Johnson-Laird, 1972). Research into content and context effects within conditional reasoning continues to the present (e.g. De Neys, Schaeken, & d'Ydewalle, 2003; Markovits & Potvin, 2001).

The research cited above clearly illustrates the influence of both content and context in the domain of syllogistic reasoning. We will now consider such influences with the Wason selection task and the conditions in which reasoning on this task is facilitated.

1.3 Facilitation Effect on the Wason Selection Task

A further task demonstrating the influence of content and context on deductive reasoning is the selection task. The selection task was introduced by Wason (1968) and has become the most widely used problem in the study of human reasoning (Evans, Newstead & Byrne, 1993). The standard abstract form of the task comprises of four hypothetical cards showing either a number or a letter and a statement, "If a card has a vowel on one side, then it has an even number on the other side"; traditionally, the cards shown denote E, K, 4, and 7. The cards are referred to as the p card (vowel), not-p card (¬p, non-vowel), q card (even number) and not-q card (¬q,

non-even number). Participants are asked, "Which cards would you need to turn over to test whether the statement is true or false?" According to the logical truth-table for material implication the only selections to falsify the rule would be the p-card (or vowel, E) and the ¬q card (non-even number, 7). If, as the logicists would have us believe, people employ formal logic in their reasoning, then this should be a relatively easy task as nothing more is required than implementing the appropriate truth-table. However, typical solution rates are below 10% (Wason & Johnson-Laird, 1972) suggesting that people are using alternative modes of reasoning when faced with the task. Most participants choose the p and q cards or the p only card. Thus, they appear to be attempting to confirm the rule, rather than disconfirm it. This led to the earliest theory of the selection task proposed by Wason himself (Wason, 1968; Johnson-Laird & Wason, 1970) known as *confirmation (or verification) bias* which assumes that the statement is considered as a biconditional rather than a material implication conditional. That is, given if p then q as true, one assumes that q then p, is also true.

However, Evans & Lynch (1973) found participants would select the items named in the statement, even when negatives were used, irrespective of the conditional rule, a phenomenon they labelled *matching bias*. Attempts have been made to simplify the task with the aim of inducing facilitation. Indeed, if the task involves only the q and ¬q instances, referred to as the Reduced Array Selection Task (RAST), facilitation is enhanced with participants making more ¬q selections (Johnson-Laird & Wason, 1970; Wason & Green, 1984). Dominowski (1995) reviews a number of variations of the selection task and the facilitatory effects these variations have on task performance.

The form of the selection task to enhance performance significantly is the thematic selection task, in which it is observed that the logically correct p and ¬q cards are chosen. Thematic versions of the task employ a range of concrete, contextually-rich scenarios that relate directly to personal experience. The first of these thematic tasks, the so-called transport rule, was devised by Wason & Shapiro (1971) in which 63% chose the logically correct solution compared to just 6% with the abstract version of the rule.

Familiarity with a rule is also an important factor as shown by Johnson-Laird, Legrenzi & Legrenzi (1972) who used an unusual postal rule that was prevalent in Britian and Hong Kong at the time. The postal rule stated that if a letter is sealed then a first class stamp must be placed on the letter. A facilitation effect was once again found with this task in comparison to the standard Wason selection task. However, the postal rule task has been criticised for being "too realistic" (Evans, 1982, p.182) as participants may be using their knowledge of postal rules and not be engaging in any form of reasoning whatsoever. This criticism has some support from the finding with American participants unfamiliar with the postal rule who showed no facilitation effect (Griggs & Cox, 1982).

Despite some early reservations the facilitation effect with the selection task is now well established as the evidence in its favour has grown. One study involving realistic material but not involving a rule that cues memory or prior knowledge is that of D'Andrade (reported in Rumelhart, 1980). D'Andrade devised the so-called *Sears* rule in which an American department store (Sears) issues the following rule to its staff, "If any purchase exceeds $30, then the receipt must have the signature of the

departmental manager on the back". With a rationale for the rule a facilitation effect was observed.

A study that has been replicated many times is that of Griggs & Cox (1982) who introduced the *Drinking Age* rule. Participants are cued to the role of a Policeman enforcing the following rule: "If a person is drinking beer, then the person must be 19 years of age" (note: 19 was the legal age to drink alcohol in Florida at the time). Four cards were presented showing Beer (p card), Coke (¬p card), 16 (¬q card) and 22 (q card). As with the indicative form of the task the p and ¬q cards were considered the correct selections and over 70% of participants chose these two cards compared to none for the standard task.

However, the concreteness of a thematic task is not sufficient by itself to produce facilitation. Gilhooly & Falconer (1974) used the following rule: " Every card which has Manchester on one side has a car on the other side" but achieved only modest improvement on the abstract selection task. In some cases facilitation may not be achieved at all. Manktelow & Evans (1979) used statements concerning food and drink such as, "If I eat fish, then I drink gin" and found performance to be no better than with an abstract version. In this study the associations between food and drink may be too arbitrary to produce facilitation.

In sum, concrete material per se does not lead to facilitation on the selection task. However, scenarios analogical to everyday experience, such as the *Sears* rule or *Drinking Age* rule or those that can be supported by a rationale, for example the *postal*

rule do improve performance significantly. Dominowski (1995) provides an in depth analysis of the factors that contribute to the facilitation effect.

Impressively, thematic versions of the reduced array selection task (RAST), in which only the q and ¬q cards are made available, have produced facilitatory effects with children (Girotto, Blaye, & Farioli, 1989; Girotto & Light, 1991; Girotto, Light, & Colbourn, 1989) and Cummins (1996a) has obtained such results with children as young as 3 - 4 years old.

1.4 Deontic Reasoning

The word *deontic* is derived from the Greek word *dei*, meaning *one must*. Deontic reasoning involves reasoning about what one must, should or ought to do. As suggested by Cummins (1996b) our social institutions presuppose the capacity to understand and reason with permissions, obligations, and prohibitions and consequences arise if deontic reasoning is not followed. Further, deontic reasoning is assumed in child-rearing and social interactions in the form of permissions, promises, warnings or threats. The next section will outline deontological ethics, a philosophical theory of moral reasoning, otherwise known as deontology.

1.5 Deontological Ethics

Moral philosophy (or ethics) is concerned with understanding the nature of human values, of how individuals ought to live, and what should be considered as correct conduct. What constitutes a good moral life has been a central question for moral philosophers for several centuries.

A distinction is made in moral philosophy between *teleological* and *deontological* theories. Teleological theories consider moral behaviour in terms of its consequences. Arguably the most influential of such theories has been *Utilitarianism* (J.S. Mill, 1861) which equates *good* consequences with happiness; hence Mill's dictum, "The greatest happiness for the greatest number". In contrast, deontological theories emphasise ones duties or obligations to perform certain actions and refrain from others, irrespective of consequences. Kant (1785/1991), who was influenced by the protestant orthodoxy of his day, stressed duties ("*duty for duty's sake*") and introduced the categorical imperative as a central concept of his ethical theory.

A further distinction to be made is between theoretical and practical reasoning, as first suggested by Aristotle. The former is reasoning in the formal sense, for example reasoning with syllogisms, while the latter involves reasoning when a decision is to be made or an action followed. The selection task offers a good example for distinguishing between these two forms of reasoning. The abstract form of the selection task, as devised by Wason (1968), requires the individual to reason *about* a rule, that is to establish its truth status, and thus involves theoretical reasoning. In contrast the *thematic* versions of the selection task involve reasoning *from* a rule; that is, considering whether a rule is being adhered to, and therefore involves practical reasoning (Manktelow, 2004).

Deontic reasoning is, therefore, a form of practical reasoning (Audi, 1989). Deontic logic attempts to formalise principles when reasoning with permissions, obligations and prohibitions. It is applicable to theoretical reasoning more readily than practical reasoning and has important differences from normative logic, the primary difference

being that deontic logic is not truth functional (see Chellas, 1980 for a review of deontic logics).

1.6 Theories of Deontic Reasoning

Five alternative theories that have attempted to explain deontic reasoning will be considered. The first is Pragmatic Reasoning Schema (PRS) theory proposed by Cheng & Holyoak (1985) in their seminal paper. An alternative evolutionary approach, Social Contract Theory (SCT) was offered by Cosmides (1989) which has sparked much debate and discussion. Manktelow & Over (1991; 1995) modified the Mental Models Theory (MMT) of Johnson-Laird & Byrne (1991) to incorporate personal preferences within deontic conditional statements. Probabilistic theories have also been offered as explanations of deontic reasoning. Oaksford & Chater (1994) adapted their influential Information Gain theory (IGT) to explain reasoning in deontic contexts. More recently, Evans, Handley & Over (2003) and Evans & Over (2004) have proposed a dual process theory of reasoning which has been applied to deontic reasoning. Each of these theories will be considered in turn.

Cheng & Holyoak (1985) offered the first theory to explain the influence of content and the facilitation effect in the Wason selection task in terms of deontic reasoning. They identified a consistent feature in improved performance with versions of the selection task; facilitation was found when the task involved deontic rules. Cheng & Holyoak (1985) proposed specialised schemas for conditional permissions and later similar schemas were proposed for obligations (Cheng, Holyoak, Nisbett, & Oliver, 1986; Politzer & Nguyen-Xuan, 1992).

In their paper Cheng & Holyoak (1985) reject two candidate theories of the time; the natural-logic view and the specific-experience view. The natural-logic view (e.g. Braine, 1978; Braine, Reiser & Rumain, 1984; Rips, 1983) developed from the inadequacies of formal logic as an explanation of reasoning (e.g. Henle, 1962; Inhelder & Piaget, 1958) suggests there are a set of *natural* inferential rules which apply across all reasoning contexts. The specific-experience view (e.g. Griggs & Cox 1982; Manktelow & Evans, 1979; Reich & Ruth, 1982) holds that facilitation on the Wason selection task is due to ones memory of domain-specific experiences (i.e. postal or drinking rules).

Cheng & Holyoak (1985) claim, however, that there are two findings that these theories cannot explain. Firstly, facilitation is not always assured when domain-specific experiences would be expected to be involved (Manktelow & Evans, 1979) and facilitation has been observed with novel scenarios (D'Andrade, reported in Rumelhart, 1980). Secondly, there are differences in the card selections made by participants; with *symbolic* or *abstract* problems card selections match the terms mentioned in the rule (matching bias), whereas with *realistic* problems participants tend to select the p & q cards, therefore verifying the rule (verification bias).

Pragmatic reasoning schemas are generalized, context-sensitive rules abstracted from personal experience of permissions and obligations. Such schemas are defined in terms of goals such as taking an appropriate action or making predictions about future events.

Cheng & Holyoak (1985) offer four production rules of the permission schema:

Rule P1: If the action is to be taken, then the precondition must be satisfied.

Rule P2: If the action is not to be taken, then the precondition need not be satisfied.

Rule P3: If the precondition is satisfied, then the action may be taken.

Rule P4: If the precondition is not satisfied, then the action must not be taken.

When faced with an arbitrary rule it is unlikely that any of the permission rules will be evoked. However, a rule that does evoke the permission schema will afford the set of rules becoming available. The rules are not meant to be equivalent to propositional logic but there is a match between each rule and the four inference patterns of the material conditional. Rule 1 is associated with Modus Ponens; Rules 2 and 3 block the fallacies of Denying the Antecedent, and Affirming the Consequent, respectively; and Rule 4 corresponds to Modus Tollens. The permission schema is context-sensitive and the rules act as heuristics rather than logical inferences. Cheng & Holyoak were able to support their theory with a series of experiments, including an unfamiliar immigration regulation and one in which participants were given an abstract permission rule, such as "If one is to take action A, then one must first satisfy precondition P"; 61% of participants solved the abstract permission problem compared to only 19% solution rate for an abstract version of the selection task.

A corresponding obligation schema was later proposed with its own set of production rules (Cheng, Holyoak, Nisbett, & Oliver, 1986; Politzer & Nguyen-Xuan, 1992):

Rule O1: If the precondition is satisfied, then the action must be taken.

Rule O2: If the precondition is not satisfied, then the action need not be taken.

Rule O3: If the action is to be taken, then the precondition may have been satisfied.

Rule O4: If the action is not to be taken, then the precondition must not have been satisfied.

Holyoak and Cheng (1995) modified their theory slightly by proposing that a conditional permission bestows a *right* to take a regulated action, while an obligation imposes a *duty* to take a particular action. The notion of rights and duties obviously iterates earlier philosophical writings. Holyoak and Cheng believe this complementarity of rights and duties to be consistent with other relationships represented in production systems, such as *parent of* and *child of*. Of course, deontic conditionals involve voluntary actions and people may violate their rights and duties. Whether rights and duties contribute to the coherency or parsimony of PRS theory is open to debate (see commentaries to Holyoak & Cheng, 1995).

The Pragmatic Reasoning Schema theory has certainly been influential but has not gone without criticism. Manktelow & Over (1991) note that true permissions are optional and should contain the word, "may", however, rules 1 and 4 of Cheng & Holyoak's (1985) permission schema contain the modal, "must", suggesting an obligation. One may also argue that PRS theory maintains the logicist perspective with its permission rules corresponding to inference patterns of the material conditional. Indeed, O'Brien (1995), a logicist himself, has endorsed a schema theory such as PRS as an essential element of a contemporary logic-based theory.

An alternative explanation to PRS theory was offered by Cosmides (1989; Cosmides & Tooby, 1989, 1992, 1994) who applied earlier ideas of evolutionary biologists (Axelrod, 1984; Axelrod & Hamilton, 1981; Trivers, 1971) to the study of human reasoning. Known as *Social Contract Theory* (SCT) or *Social Exchange Theory* (SET), Cosmides proposes that our evolutionary past has provided humans with innate *cheater-detection algorithms* (also referred to as *Darwinian algorithms*) which are evoked in situations of social exchange and operate according to *cost-benefit* principles. More specifically, the belief is that because hominids have spent 99% of their evolutionary history as hunter-gatherers, a practice developed in the Pleistocene era, they have adapted to engage in social exchange, which Cosmides (1989, p.196) defines as, "cooperation between two or more individuals for mutual benefit". Few species have evolved the capacity for social exchange, however, it is this which is considered essential for cooperative social living. Thus, humans have evolved, "Darwinian algorithms – specialized learning mechanisms that organize experience into adaptively meaningful schemas or frames" (Cosmides, 1989, p.196). Once activated, the adaptive schema, "focus attention, organize perception and memory, and call up specialized procedural knowledge that will lead to domain-appropriate inferences, judgements and choices" (Cosmides, 1989, p.196).

Social contract theory challenges the view that the mind comprises of relatively few domain-general mechanisms and argues for a greater number of special-purpose modules, such as cheater-detection algorithms. Cosmides also challenged Cheng & Holyoak (1985) on methodological grounds, claiming that their findings could be interpreted in terms of Tversky & Kahneman's (1973) *availability heuristic* theory.

Cosmides (1989) reinterpreted the Wason selection task in terms of a cost-benefit structure involving the testing of a social contract rule. A standard social contract would take the form, "If you take the benefit, then you pay the cost". In this instance the logically correct answer for card selections would correspond with the social contract answer when seeking to detect a violator, that is, the p and not-q cards (benefit accepted and cost not paid, respectively). Ingeniously, Cosmides included a switched social contract of the form, "If you pay the cost, then you take the benefit". In this case the logically correct choice is again the p and not-q cards, however, the predicted response according to a social contract interpretation is the not-p and q card choices (cost not paid and benefit taken, respectively). Further, Cosmides used rules that would be both familiar and unfamiliar to participants. Over 70% of participants chose cards that supported the social contract theory with familiar and unfamiliar materials, including the predicted selections for a switched social contract rule. Hence, social contract theory has become a leading explanation for both the Wason selection task and for human reasoning generally. It has gained support from some quarters (e.g. Gigerenzer & Hug, 1992) and raises important questions about what evolutionary advantage there is in being able to reason and why and how such a cognitive capacity evolved in the first place. However, the theory has not gone unchallenged (see for example, Cheng & Holyoak, 1989). A major difficulty for social contract theory is explaining deontic reasoning in non-adaptive contexts. Manktelow & Over (1990) devised a precautionary rule, "If you clear up spilt blood, then you must wear rubber gloves" which produced a clear facilitatory effect. However, it is difficult to interpret such a rule as a cost-benefit relationship. In response, Cosmides & Tooby (1992) suggested Manktelow & Over had identified

another evolutionary module involving precautionary rules; a view consistent with the massively modular perspective of the human mind.

The most influential theory of reasoning in the past decade has arguably been Mental Models theory (MMT). Originally applied to the comprehension of discourse and to syllogistic reasoning (Johnson-Laird, 1975, 1983) mental models theory was more fully developed as a comprehensive theory of deduction by Johnson-Laird & Byrne (1991). Mental models theory is based on the earlier work of Craik (1943) who postulated that the mind derives small-scale models of external reality. A key difference between the two theorists is that Craik considers models as parallels or imitates of reality but not necessarily having the same structure as what they represent, whereas Johnson-Laird believes they are iconic, that is, the mental model is a mirror of what it represents. Johnson-Laird & Byrne (1991) proposed three cognitive stages to their theory comprising of stage 1 in which a mental model of the given premises is derived using ones general knowledge and knowledge of language, stage 2 where reasoners attempt to assert a conclusion from the given premises, and stage 3 in which alternative models (counter-examples) are sought in order to test a putative conclusion and if no credible alternative can be found the conclusion is proposed as being valid. Johnson-Laird & Byrne (1991) presented the mental models theory as an explanation of the major areas of deduction, including propositional reasoning, relational reasoning, and quantificational reasoning. In the decade and a half since their seminal monograph, Johnson-Laird and his many collaborators have used mental models theory to explain many other areas of thinking and reasoning, including spatial reasoning (Byrne & Johnson-Laird, 1989), reasoning with conjunctives and disjunctives (Garcia-Madruga, Moreno, Carriedo, Gutierrez &

Johnson-Laird, 2001), propositional reasoning (Johnson-Laird, Byrne & Schaeken, 1992) and extensional reasoning (Johnson-Laird, Legrenzi, Girotto, Legrenzi & Caverni, 1999).

The mental models account of the selection task (Johnson-Laird & Byrne, 1991) is a modification of an earlier proposal by Johnson-Laird & Wason (1970) in which models replace truth tables. Presented with the cards in the selection task, participants are said to consider those cards that are explicitly represented in their models of the rule and they then select those cards for which the hidden value could have a bearing on the truth or falsity of the rule. For instance, given the rule "If there is an A on one side of a card then there is a 2 on the other side" and four cards: A, B, 2, 3 the interpretation of the rule will typically determine the selection of cards. If the rule is interpreted as a conditional then the following model will be derived:

    [A]   2

        …

The [A] indicates that A has been exhaustively represented (i.e. A cannot occur in any other model), while 2 indicates a non-exhaustive representation. The ellipsis (…) denotes implicit information concerning alternative models. A conditional interpretation will result in only the A card being selected (although both the A and 2 cards will be considered) because it alone has a hidden value that could bear on the truth or falsity of the rule. A bi-conditional interpretation of the rule yields the following model:

    [A]   [2]

        …

Both the [A] and [2] are exhaustively represented but alternative models are possible if made explicit. With a bi-conditional interpretation both the A and 2 cards will be selected according to mental model theory. Only when the alternative models are explicitly fleshed out will the negative consequent, required to produce the logically correct solution, be represented.

The mental model theory offers a parsimonious account of conditional rules used in some deontic contexts. The unfamiliar rule used by Cosmides (1989), "If a man has a tattoo on his face then he eats cassava root" given within a vignette about an anthropologist is interpreted by participants as meaning, "A man may eat cassava root only if he has a tattoo" according to model theory (Johnson-Laird & Byrne, 1991, p. 78). This re-interpretation yields the following models:

[eating cassava]   tattoo

¬ eating cassava   [¬ tattoo]

…

Here two models are derived due to ones general knowledge of permissions. The first model represents the interpretation of the given statement according to model theory where [eating cassava] is exhaustively represented and tattoo is not. The second model is the counterfactual where ¬ means not or negation and is intimated by our knowledge of permissions and obligations.

In his more recent writings Johnson-Laird has presented mental models as only one of several strategies that individuals may employ when reasoning (Johnson-Laird, 2005; Johnson-Laird, Savary & Bucciarelli, 2000).

A limitation of mental model theory is its inability to deal with content and context effects (Bonatti, 1994; O'Brien, Braine, & Yang, 1994). Experimental studies conducted by Johnson-Laird and colleagues have often used abstract materials, devoid of *real world* knowledge. This has lead some to claim that at a deeper level mental models are no more than formal logical rules (Stenning & Oaksford, 1993). The theory has also been questioned in its handling of bias effects (Evans & Over, 1996a).

A modified version of mental models theory proposed to explain deontic reasoning is that of Manktelow & Over (1991; 1995). Their theory is concerned with *practical* deontic reasoning (reasoning about actions) rather than *theoretical* deontic reasoning (reasoning about states of affairs). They start with the premise that one must ask why individuals reason or act as they do and to answer this Manktelow & Over adopt terms from decision theory – subjective utility and probability and benefits and costs (features not included in the original mental models theory). In any social dialogue or interaction one must consider personal preferences (subjective utilities) which may be influenced by moral, social or prudential reasons. Using an experimental paradigm of alternate perspectives (the experimental procedure is detailed in section 2.3 below) they were able to illustrate the benefits and costs of an outcome between an agent and actor.

The theory of Manktelow & Over is a semantic one, in comparison with the schema theories of Cheng & Holyoak (1985) and Cosmides (1989) which are syntactic. However, Manktelow & Over do not completely reject schema theories and believe that a hybrid theory consisting of schemas and mental models may be possible. Mental models are the basic unit of their theory with mental models being devised

from pragmatic information as well as the semantics of the rules. Each party (agent or actor) will represent a different set of mental models based on the benefits and costs of the possible outcomes. For example, in an agent-actor dialogue, the agent may utter, "If you p, then you may q". It is assumed that the actor must prefer q to not-q and the agent must prefer p to not-p. This state of affairs can be represented, thus:

Agent's Perspective   *Actor's Perspective*

p+   q   +   p   q+   +

…   …

The p+ denotes a benefit for the agent and q+ a benefit for the actor. The + to the right shows the overall utility for that particular outcome; the agent prefers outcomes where p is fulfilled, the actor prefers outcomes where q is fulfilled. The three ellipsis indicate implicit models that will only be fleshed out (made explicit) if the task demands it. A cost, denoted by a negative sign (-), arises when one is *cheated*, from having one's expectation unfulfilled or one's authority defied. Thus, all mental models can be made explicit from both the agents' and actor's perspectives and their personal preferences defined. Given a deontic conditional statement, four outcomes are possible from each perspective, two positive (benefiting the actor &/or agent) and two negative (representing a cost to the actor &/or agent); in fact, only one scenario will result in both the actor and agent benefiting and that is when the p and q components of the statement are both made true (as shown in the example above).

A key advantage of this approach is that the inclusion of personal preferences suggests why an individual would make a deontic statement in the first place. Further,

the approach makes a number of novel predictions, some of which have yet to be tested empirically. However, the theory has not subsequently been developed by Manktelow & Over or their collaborators and therefore remains speculative in a number of regards. For example, Manktelow & Over (1995) indicate that additional pragmatic factors could be incorporated in one's reasoning by the addition of further columns representing benefits and costs. How this would work in practice and what effects the complexity of models would have on one's reasoning is not explained.

The fourth type of theory to be discussed in this section is dual process theories. Dual process theories of reasoning have a long history dating back to Aristotle. More recent proposals have been developed independently, yet share a number of similar features. Sloman (1996) proposed two systems of reasoning; associative and rule-based. The associative reasoning component is taken from James (1890) and is based on associative parallel processing systems (see McClelland & Rumelhart, 1986). The computational task of the associative system reflects the similarity and temporal structure of stimuli. The second system has rules as the basic form of representation which comprises of a number of guiding principles. Rules are abstractions that apply to all statements, having a well-specified, symbolic structure. There are different kinds of rules such as computer programmes, laws and rules of logic.

Evans has been a key proponent of dual process theory for many years. Evans (1984; 1989) proposed the heuristic-analytic theory to explain the common errors and biases people made when undertaking conditional reasoning tasks. The heuristic stage is described as *pre-attentive* and involves the unconscious process of selecting relevant information from the source material and from memory. It is the processing at this

30

stage which can lead to errors as participants are likely to be influenced by the salience of given information, linguistic factors and the effects of prior knowledge (Evans, 1984). Only if analytical processes are applied to the representation of a reasoning problem will logical competence be observed (Evans, 1989). Evans did not develop the analytic stage of his theory and the early emphasis was on the heuristic stage and its application to human bias in reasoning.

The greatest influence on Evans and colleagues more recent theorising has been the research into individual differences in reasoning. Stanovich (1999) provided a monograph of individual differences in human reasoning which chronicled much of the work of Stanovich & West (1997; 1998). Stanovich & West (1997; 1998) took a novel approach to investigating reasoning by focusing on the minority of individuals who were able to produce logically correct solutions to the Wason selection task and other reasoning problems. They found such individuals to possess a logical competence that eluded the majority. Two systems were proposed by Stanovich (1999) to explain the differences in reasoning observed between individuals which were termed System 1 and System 2. The two systems are analogous to Sloman's (1996) associative – rule-based theory and Evans' (1984; 1989) heuristic-analytic theory. It is suggested that both systems have evolved with System 1 being observed in the brains of all higher mammals but only System 2 is to be found in humans. System 1 processing involves simple learning, such as conditioning, is influenced by heuristics and pragmatic knowledge and by content and context. Thus, belief-biased effects can be explained in terms of System 1 thinking. Inferences made using System 1 tend to be drawn rapidly and automatically without conscious reflection. Inferences made using System 1 do not necessarily lead to logical errors, however, the solution is

likely to have been drawn on the basis of high probability or plausibility, rather than by logical reasoning. Stanovich (1999) proposes that System 2 evolved after System 1 and is linked to language and involves reflective consciousness. Processing is explicit, conscious and analytic and operates via verbal working memory. System 2 is much slower than System 1 as processing is sequential. Stanovich (1999) and Stanovich & West (2000) provide evidence to suggest that System 2 processing correlates with intelligence, with those individuals having a high IQ being less prone to reasoning biases and more likely to make logically correct inferences. System 2 is a general human ability but it appears that some individuals have more of it than others.

Evans & Over (1996b) proposed a similar theory of rationality in which the key processes were Rationality 1 and Rationality 2. This was a development of Evans' previous work, with the two forms of rationality being equivalent to heuristic and analytic processes, respectively. In their most recent writings Evans, Over and their colleagues have adopted the System 1 / System 2 distinction into their theorising and have begun to develop dual processing theory in such areas as hypothetical thinking (Evans, Over & Handley, 2003; Evans & Over, 2004), causal conditional reasoning (Over, Hadjichristidis, Evans, Handley & Sloman, 2007) and deontic reasoning (Over, Manktelow & Hadjichristidis, 2004). Aspects of this work will be reviewed below.

Finally, probabilistic theories of deontic reasoning will be considered. Oaksford & Chater (1994) reviewed the vast data from studies investigating the selection task and observed that the patterns of card choices could be predicted using a Bayesian analysis. This approach is influenced by J.R. Anderson's (1990) *rational analysis* which distinguishes between *normative* and *adaptive* rationality. The former is

associated with rules of mathematics and logic, while the latter considers rationality as an organism's adaptation to the environment; the rational analysis approach is concerned with adaptive rationality.

According to Oaksford & Chater the selection task is approached by participants as a task involving gaining information by reducing uncertainty rather than as a logical reasoning task. Several assumptions are incorporated into this model, including the *rarity assumption,* which assumes that properties in causal relations are rare in the environment. Information gain is, therefore, inversely related to probability; the lower the probability of an event the more informative that event is. It is proposed that, when faced with the selection task, participants adopt one of two alternative hypotheses termed the *dependence model* (in which p depends on q) and the *independence model* (in which p and q are independent). The likelihood or prior probability for each hypothesis is taken to be 0.5. The hypothetical task of examining each card provides data allowing one to update their hypothesis (information gain or Ig). However, as the task is hypothetical in that participants never actually turn any cards, all possible outcomes must be considered; hence the measure taken is expected information gain E(Ig). A noise or error factor is included, resulting in the calculation of scaled expected information gain or SE(Ig) for each individual card. The combination of these factors produces scaled expected information gain's in the ordering of p > q > ¬q > ¬p, which corresponds to the frequency of card selections with the abstract version of the selection task.

For deontic versions of the selection task Oaksford & Chater (1994) refocus their probabilistic model away from *rule testing* and on to *rule use.* The dependence and

independence models are used to calculate expected utilities and participants are believed to use the rules to maximise expected utility. A small fixed cost is assumed for turning any card. Also, participants associate particular utilities with particular card combinations, depending on perspective and the rule. Oaksford and Chater use the terms enforcer and actor when referring to perspective effects. The enforcer's goal is to discover rule violations, that is, where the actor performs the action without satisfying the condition. This is modelled by assigning a positive utility to instances of rule violation that is larger than the cost of turning over a card. The actor's goal is to discover instances of unfairness, where the enforcer disallows the action even though the actor satisfies the condition. This is modelled by assigning a positive utility to uncovering instances of unfairness that is larger than the cost of turning a card. For a permission rule utility is incorporated in to the probability model by assigning greater weights to the not-condition (¬p) and action (q) card combination for the enforcer, and the condition (p) and not-action (¬q) card combinations for the actor. The probability of the independence model, P(MI), in the abstract task represents the probability of p and q being independent but in the thematic task P(MI) represents the probability that an individual is disobeying the rule. The card selections are modelled in the same way as for the abstract task except that expected utilities for each card is calculated and the rarity assumption is assumed not to hold for deontic rules. Card selections are indicated in the model by the greatest expected utility scores. With a permission rule, the model selects the not-condition (¬p) and the action card (q) for the enforcer perspective, and the condition (p) and not-action (¬q) cards for the perspective of the actor. Taking the findings from experiments involving deontic conditionals by Cheng & Holyoak (1985), Cosmides (1989), Gigerenzer & Hug (1992) and Manktelow & Over (1991) it has been found that, "(t)hese predictions

agree perfectly with the results of the studies indicated" (Oaksford & Chater, 1994, p.623).

A number of criticisms have been levelled at Information Gain theory (see the commentaries of Almor & Sloman, 1996; Laming, 1996; Evans & Over, 1996a). Possibly, the most damning criticism is that the theory offers a statistical theory of reasoning but fails to provide a psychological explanation of the underlying cognitive processes (Evans & Over, 1996a; 1996b). Furthermore, how accurately any model of this sort can model performance is debatable. Girotto (1995) considers in detail the card selections made in a number of deontic tasks illustrating perspective effects and argues persuasively that performance is influenced by the rationale given to the participant with consistency of findings being far from perfect.

Oaksford and Chater's probabilistic approach to reasoning has been influential in the development of the recent probabilistic theory of Evans and colleagues (Evans, Handley, & Over, 2003; Evans, Over, & Handley, 2003; Evans & Over, 2004).

In their theory of hypothetical thinking, Evans, Over & Handley (2003) have proposed three principles:

Singularity: only one hypothetical possibility, or mental model may be considered at any one time;

Relevance: the most relevant (most plausible or probable) model is considered in the given context;

Satisficing: models are evaluated with reference to the current goal and accepted if satisfactory.

Evidence for the singularity principle comes from studies involving hypothesis-generating tasks in which it appears that participants hold but one hypothesis. For example, Bruner, Goodnow & Austin's (1956) found children could only consider a single hypothesis when completing a concept learning task. Further, Wason's (1960) finding with the "2, 4, 6" problem illustrated the difficulty individuals have in considering more than one hypothesis when undertaking a relatively simple cognitive task. The difficulty of reasoning with disjunctives is also taken as evidence for the singularity principle. A classic demonstration of this difficulty is Wason's THOG problem (Wason & Brooks, 1979).

Three factors determine the construction of the most plausible model, according to the relevance principle. These factors include the features of the task or environment, the current *goal* of the individual, and background knowledge. The principle is clearly influenced by the relevance theory of linguistic pragmatics and discourse comprehension developed by Sperber & Wilson (1986; 1995). Selecting what is perceived as the most relevant information can be manipulated by the instructions given while participants undertake a task. For example, with the truth table task false possibilities can be evoked by instructing participants to focus on particular cases (Evans, Newstead & Byrne, 1993; Evans, Legrenzi & Girotto, 1999). Thematic versions of the Wason selection task are further examples of how relevance is drawn from the content and context of the task. For example, studies of perspective effects with thematic selection tasks demonstrate how opposing roles may produce distinctive relevance.

The third factor, the *satisficing* principle, is taken from the writings of Simon (1982). The idea is that due to our limited information processing capacity we are unable to consider all possible options when making a rational choice or decision. Therefore, we are likely to opt for a *satisfactory* choice via the use of a suitable heuristic. According to Evans et al. satisficing occurs at the stage of explicitly evaluating models. The single model is considered as a suitable candidate for satisficing the current goal. If satisfactory, the model is accepted, if not it is rejected and an alternative is generated and evaluated. Evans et al. reject the notion of optimisation in normative decision-making theory and propose that decisions are most likely to be made using the satisficing principle, as Simon (1982) suggested. They also believe that reasoning is by default probabilistic and pure deductive reasoning extremely rare.

The probabilistic element of their theorising has been developed in their *suppositional theory*, a form of hypothetical thinking (Evans, Handley & Over, 2003; Evans & Over, 2004). It is proposed that a conditional statement of the form, If p, then q is interpreted as the probability of the consequent, given the antecedent (P(q|p)). This hypothesis is called the *conditional probability hypothesis* and is a view held by a number of philosophical logicians (see Bennett, 2003). An example from the Evans et al. paper will illustrate the conditional probability hypothesis. Take the statement:

If the car's battery is dead then it will not start

Knowing that the car's battery is dead suggests a high probability that the car will not start and thus the P(q|p) is high.

The conditional probability hypothesis is one assumption of Evans et al.'s probability theory. A second assumption is that premises can be extended by pragmatic implicatures, that is, inferences may be invited by the content and/or context of an utterance. Evans & Over (2004) give the following statement as an example:

If you clean my car then you can borrow it tonight

By pragmatic implication this statement is biconditional; if you do not clean the car then you cannot borrow it.

A third assumption involves System 1 / System 2 thinking. System 1 has limited reasoning powers and a conditional statement, if consistent with one's beliefs, would be taken to mean the subjective value of $P(q|p)$. Only System 2 thinking adheres to the principles of deductive reasoning.

The most influential theories of deontic reasoning, over the past two decades, have been reviewed in this section. The schematic theories of Cheng & Holyoak (1985), with their Pragmatic Reasoning Schema (PRS) theory, and Cosmides' (1989) Social Contract Theory (SCT) sparked the theorising of the thematic selection task and introduced deontic reasoning to psychology. The Mental Models Theory (MMT) of Johnson-Laird & Byrne (1991) is a more general theory of reasoning that has been particularly influential; modifications of the theory have also been proposed to explain deontic reasoning (Manktelow & Over, 1995). Dual process theories have re-emerged, offering a more complete explanation of two cognitive systems; System 1 reliant on reasoning with probabilities and System 2 that can reason logically (Stanovich, 1999; Evans & Over, 1996b). More recently, probabilistic theories of

reasoning have been proposed that have switched the emphasis from trying to explain logical competence (or lack of) to considering the interpretation of tasks as one involving possibilities or probabilities. Oaksford & Chater's (1994) theory of Information Gain has been one such theory that has gained in stature as it extends its range of applications, including an explanation of deontic reasoning. An alternative probability theory has emerged from the work of Evans and colleagues into hypothetical thinking (Evans, Handley & Over, 2003; Evans, Over & Handley, 2003; Evans & Over, 2004). There are, therefore, several candidate theories of deontic reasoning.

The approach taken in this thesis is to adopt the probability theory of Evans and colleagues and to consider pragmatic factors of deontic reasoning as residing in System 1 thinking. Reasoning with pragmatic content typically involves satisficing and reducing cognitive load by using only relevant information. It is also proposed that judgements involving pragmatics are determined by the conditional probability hypothesis. Furthermore, this probabilistic approach will be applied to experiments involving judgement revision and conjunctive and disjunctive statements.

Chapter 2

Pragmatics and Decision-Making

2.1 Introduction

Pragmatics has traditionally been the domain of linguists but is now having an increasing role in the study of reasoning. The relationship between the pragmatics of language and that of reasoning will be outlined before discussing specific pragmatic factors that have been found to influence deontic reasoning to date. Decision-making is another field that has traditionally been investigated separately from reasoning. However, there has been a gradual realisation that many reasoning tasks involve decision-making and situations requiring a decision or judgement include a reasoning element. Theories of decision-making will be described and an argument made for there being similar processes involved in reasoning and decision-making within the deontic domain.

2.2 Pragmatics of Language

The term *pragmatics* was introduced by Morris (1938) who also made the distinction between syntax, semantics and pragmatics. Mey (1993, p.3) defines pragmatics, "as the science of language seen in relation to its users … as it is used by real, live people, for their own purposes and within their limitations and affordances". Although many definitions of the term exist (see Levinson, 1983) most explanations agree that pragmatics, within a linguistic or reasoning context, involves acquired *world knowledge* in which the content of discourse may yield its effects.

Possibly the most widely cited theory of pragmatics is that of Grice (1975) who

showed that formal logic does not correspond to natural language and yet everyday language provides statements that are just as precise as logic and are understood in a given context. In his theory of *conversational implicatures* Grice proposes maxims or principles which guide the meaning of spoken language. His maxims include, quantity (provide only sufficient information), quality (be truthful), relation (be relevant) and manner (be clear). There is experimental evidence for the use of these maxims (Newstead, 1989; 1995) and Evans & Over (2004) in their suppositional theory incorporated Gricean maxims in what they refer to as pragmatic implicatures. The research reported in this thesis assumes Grice's *principle of co-operation* as integral between speaker and addressee.

2.3 Pragmatic Effects of Deontic Reasoning

Increasingly, the phrase pragmatics is being used within the domain of thinking and reasoning. Cheng & Holyoak (1985) were the first to refer to pragmatic effects in relation to deontic reasoning, suggesting the range of influences on the thinking of permissions and obligations is much greater than hitherto realised. Pragmatics have been included in models of analogical reasoning (Holyoak & Thagard, 1989; Keane, 1990) and there is a growing body of research into the pragmatics of order effects (Feeney, Evans & Clibbens, 1996; Girotto, Mazzocco & Tasso, 1997; van der Henst, 1999). A recent influential theory attempting to explain pragmatics is *Relevance* theory (Sperber & Wilson, 1986, 1995) which has been applied to various pragmatic influences within thinking and reasoning (eg. Sperber, Cara & Girotto, 1995; Girotto, Mazzocco & Tasso, 1997; van der Henst, 1999).

Relevance theory (Sperber & Wilson, 1986, 1995) is an explanation of language

understanding but has been applied to reasoning. The theory develops Grice's maxims and holds that a speaker's utterance will always be relevant, even when statements appear to be unrelated. Manktelow (1999) offers the following example, "I'm going to watch TV tonight. You can borrow the car". There may be a number of reasons why the second statement is relevant to the first; obviously the context and the relationship between speaker and addressee would provide the answer. Sperber and Wilson define relevance in terms of cognitive *effect* and *effort*; the greater the relevance of a piece of information the less cognitive effort is required. Thus, an inverse relationship is observed between these cognitive components. The theory has applications – Sperber, Cara, & Girotto (1995) used it as an explanation of the selection task; and Evans & Over (1996a) proposed their heuristic stage of reasoning (System 1) involved relevance. Relevance theory has inspired dialogue between researchers in linguistics and reasoning and has raised the profile of pragmatics as an important part of everyday reasoning, a crucial component of this thesis.

Two pragmatic factors that have been found to be associated with reasoning in deontic contexts are perspective and probability. Perspective has been extensively studied usually within a deontic context of a selection task. The general finding is that when a participant is asked to play a role or take a particular perspective facilitation of predicted card selections is observed. Further, *correct* card selections can be manipulated by perspective and scenario. For instance, Manktelow & Over (1991) introduced the rule, *if you tidy your room, then you may go out to play,* which is spoken by a mother to her son. Participants are cued to the role of either the mother or the son and asked to detect any violation of this rule. Four cards are presented, *room tidied* (p card), *room not tidied* (not-p card), *went out to play* (q card), and *did not go*

out to play (not-q card). It is assumed that mother and son have personal preferences, that is, the mother would prefer a tidy room while the son would prefer to play out. Thus, those seeking a violation from the mother's perspective are more likely to choose the not-p card (*room not tidied*) and the q card (*went out to play*), while participants taking the son's perspective choose the p card (*tidied room*) and the not-q card (*did not go out to play*). Gigerenzer & Hug (1992) and to a lesser degree Politzer & Nguyen-Xuan (1992) have also demonstrated perspective effects with non-normative, yet predictable selection card choices.

The probability of violation has also been found to influence selection card choices. Kirby (1994) modified Griggs & Cox's (1982) drinking-age rule, *If a person is drinking beer, then they must be over 21 years of age*, by having potential under-age drinkers of 4, 12 and 19 years. As predicted, participants were more likely to inspect the 19 year old than the 4 year old. Manktelow, Sutherland & Over (1995) introduced the Large Array Selection Task (LAST) which is ideal for the study of probability as it allows the use of several p, not-p, q and not-q cards. With the use of the immigration rule (Cheng & Holyoak,1985) it was found that travellers from the tropics were more likely to be *checked* for appropriate vaccinations than European travellers. It, therefore, appears that situations are assessed for their probability of outcome, something that was ignored with early selection task studies.

Other research has shown that logically correct responses can be suppressed. Byrne (1989) used a number of scenarios, including the conditional, *If she has an essay to write, then she studies late in the library*. The rule was then modified by an additional statement such as, *If the library stays open, then she will study late in the library*. This

additional information facilitates or suppresses inferential reasoning, including the fallacies, denial of the antecedent (DA) and affirmation of the consequent (AC). Similar research was conducted by Stevenson & Over (1995) who introduced such conditionals as, *If John catches a fish, he will have a fish supper*. Additional statements were given, similar to Byrne's e.g. *If John catches a fish, he will have a fish supper*. Probability was manipulated in further studies by the introduction of a qualifier, for example, *John (always/ almost always/ sometimes/ rarely) catches a fish*. The findings of these two studies were similar and as predicted, but the explanations differed - Byrne preferring the idea of *suppression* of a logical inference while Stevenson & Over used the notion of *uncertainty* i.e. the probability of the truth of a statement may be reduced by additional or probabilistic information.

The current research programme extends the range of pragmatic factors influencing deontic reasoning and holds that reasoning, and thinking generally, is greatly affected by context and one's world knowledge within a given context.

2.4 Power of Source

One extension to the pragmatic factors within this thesis is the power or authority of a message source. Power has traditionally been investigated by social psychologists within the field of social cognition. However, the study of power of source within a deontic reasoning context affords the opportunity to combine these traditionally disparate fields.

The seminal work of Raven (French & Raven, 1959; Raven, 1965) identified six bases of power – reward, coercive, informational, expert, legitimate, and referent.

Such categories are not mutually exclusive when applied to an authority role as many high ranking individuals will exert most, if not all six types of influence. The research reported in this thesis involves a number of Raven's power types within individuals of varying degrees of power across a range of scenarios.

The emergence of power hierarchies in groups of unacquainted individuals has been investigated by monitoring conversations (Ng, Bell & Brook, 1993). This linguistic approach found those who spoke more and had a greater number of speaking turns were accorded a higher social ranking by the unacquainted group.

One perspective of social power proposed by social cognition theorists is a power-based schema theory (Rusk & Russell, 1988; Sande, Ellard & Ross, 1986) in which individuals pay particular attention to those who control one's outcomes (i.e.utilities) (Erber & Fiske, 1984; Neuberg & Fiske, 1987; Ruscher & Fiske, 1990). This approach considers power as hierarchical with well-developed schemas for those holding most power and less-developed schemas for subordinates who have little control over outcomes (Rodin, 1987). A recent theory explaining deontic reasoning also considers hierarchical power. The theory of *dominance hierarchies* (Cummins, 1996a, 1998, 2000) is an evolutionary theory suggesting individuals of high social ranking are more likely to influence and control the behaviour of others with violations expected of those from a low social rank.

2.5 Reasoning with Inducements

Everyday discourse involves the use of inducements. Inducements include warnings, threats and promises which can be phrased as conditional sentences, conjunctive sentences and disjunctive sentences. When an inducement is made one must consider the felicity of the statement and judge the likelihood of action that may follow. Such considerations are an essential component of deontic reasoning. Fillenbaum (1976) suggests that inducements are *purposive-causal*. That is, a conditional statement of the form, "If p, then q" consists of q being offered on my part to get something done or not done with regard to p on your part. There are, of course, important differences between the forms of inducement. Most warnings are hypothetical of the form, "If you do/do not do X, then Y will happen" (Searle, 1969). Preferences and utilities may apply to promises; the promiser would prefer an act being done to not being done and the promisee would gain from what is being offered (Searle, 1969). Of course the concept of utilities is also an established component of deontic reasoning (Manktelow & Over, 1991; Oaksford & Chater, 1994). Searle (1969) also distinguished between sincere and insincere promises. An insincere promise is one in which the speaker does not intend to undertake the act promised. The detection of an insincere promise would constitute a violator-detection as proposed by social contract theory (Cosmides, 1989). Individuals of low power are more likely to make insincere statements, that is, inducements they are unable to keep and this is one aspect of power relations investigated in this thesis.

Fillenbaum (1975, 1976) was a pioneer of research with realistic inducements. The emphasis of his research was on the interpretation of conditional inducements with the finding that promises were more likely to be interpreted as a conjunctive, while

threats tended to be interpreted as a disjunctive. However, preference of interpretation does not result in the meaning of a given statement being lost. A threat will be recognized as such whether stated as a disjunctive or conjunctive. Evidence for this comes from Springston & Clark (1973) who considered *and* and *or* statements as *pseudoimperatives*, demonstrating how statements can be paraphrased. Possibly their best known example is of a threat, "Sit down or I'll scream" which can be paraphrased as, "I'll scream if you don't sit down". Inducements phrased as conditionals, conjunctives and disjunctives are investigated within this thesis to see if they have same illocutionary force as the work of Springston & Clark (1973) would suggest.

A key advantage of studying inducements is that they tend to be more realistic and familiar compared to the traditional abstract materials and unfamiliar tasks frequently used in studying reasoning. More recent studies investigating inducements have emphasized this point (Evans & Twyman-Musgrove, 1998; Newstead, Ellis, Evans & Dennis, 1997; Newstead, Griggs & Chrostowski, 1984). The materials in the experiments of this thesis comprise of realistic, everyday statements embedded within familiar, highly contextual scenarios.

2.6 Decision-Making Theories

The normative theory of decision-making has been Subjective Expected Utility (SEU) theory which assumes that decision makers have clearly defined goals with all relevant information and evidence available to them. The aim of decision-making is to maximise one's goals (or utility), which is achieved by making a trade-off between the probability of an outcome and its expected utility.

SEU theory was introduced to psychology by von Neumann & Morgenstern (1944, 1947) but the components of the theory have a long history. The first component is probability, which did not exist as a concept before 1660 (Hacking, 1975). Pascal (1670/1941) conducted the first calculations of probability and intimated at subjective utility, the second component of SEU theory, with the so-called *Pascal's wager*; Pascal argued that the utility of believing in God and eternal life is greater than not following a religious life (interestingly, Pascal was offering a form of practical reasoning in suggesting how one *ought to* live one's life). The notion of subjective utility was further developed by Bernoulli (1738/1954) who observed that subjective utility (or value) does not directly correspond to monetary value but rather the relationship is concave. Thus, the subjective value of a monetary gain is relative according to one's current wealth.

Beach (1997) makes the distinction between paradigms, theories and models which can help unravel the plethora of research studies and multitude of explanations in the decision making field. The dominant paradigm in decision making research is *utilitarianism*, the tenet that behaviour is motivated by the attainment of valued outcomes or goals. At the next level are normative theories, based on probabilities of events and subjective utilities. At the lowest level of this hierarchy are numerous *normative models* which attempt to capture decision making behaviour. Generic terms such as (behavioural) decision theory (Einhorn & Hogarth, 1981; Slovic, Fischhoff & Lichtenstein, 1977) and (subjective) expected utility theory have encompassed a variety of normative models.

Normative models of decision-making include *Subjective Expected Utility* (SEU) theory (Savage, 1954), *Expected Utility* theory (Edwards, 1954) and *Multiattribute Utility* theory (MAUT) (Edwards & Newman, 1982; von Winterfeldt & Edwards, 1986). Descriptive models of probability, which aim to be a more accurate representation of human decision-making, have been devised and labelled, *Rank-dependent Utility* theories (Lopes, 1996; Quiggin, 1982). Prescriptive models have also been developed, including *Decision Analysis* (or *Cost-benefit Analysis*) (Brown, Kahr & Peterson, 1974; Behn & Vaupel, 1982; von Winterfeldt & Edwards, 1986) and *Cost-effectiveness Analysis* (Pliskin, Shepard & Weinstein, 1980).

The computations required for calculating utilities and probabilities are beyond our cognitive capacities (Manktelow, 1999) and, therefore, human participants are found wanting when tested on these normative concepts (Kahneman & Tversky, 1972). Thus, SEU theory, and its various forms, is an inadequate account of human decision-making from the perspective of the psychologist. The theory is said to be incomplete (Hastie, 2001) because the emphasis is on processes of inference rather than processes involving search. Theories have developed to the extent that clearly formulated axioms have been devised (de Finetti, 1937; Krantz, Luce, Suppes & Tversky, 1971; Ramsey, 1931; Savage, 1954; von Neumann & Morgenstern, 1947) allowing detailed analysis of specific aspects of SEU theory. However, systematic investigation of probability and utility has shown that human participants regularly violate these axioms.

Tversky (1969) demonstrated violation of the *transitivity* axiom using the scenario of University applicants. The axiom proposes that given a series of choices to an

individual, there should be a *weak ordering*, that is if option A is preferable to option B and option B is preferable to option C, then transitivity should be observed with option A being preferred to option C. Tversky found preference reversals with small differences between options being ignored or *underweighed* resulting in intransitivity of decision-making.

Savage (1954) formulated the *sure-thing* principle which proposed that if option 1 is preferred to option 2 in situations of certainty, then option 1 should be preferred to option 2 when in a state of uncertainty. Violations of this axiom have been demonstrated when the uncertainty concerned having passed or failed an examination (Shafir, 1993; Tversky & Shafir, 1992) and in a number of other decision-making contexts (Shafir & Tversky, 1992).

In sum, human reasoning lacks the cognitive ability to match the demands of a normative theory, such as SEU theory. When investigated experimentally, the axioms derived from normative theory are clearly violated. This suggests a theory of decision-making is required that reflects the judgements people make in real world contexts. An alternative model of decision-making will now be discussed that has greater relevance to the field of human reasoning.

A related component of decision-making is quantitative judgement which involves evaluating the options prior to making a final decision. Quantitative judgements consist of 'rating' (assigning numbers or grades), 'ranking' (placing people or objects in order on some dimension), and 'classifying' (assigning people or objects to a group). Typically, judgement experiments involve providing participants with the

attributes of options from which they evaluate; thus, judgement tasks are viewed as measures of utility. Theories of quantitative judgement tend to be based on the statistical concept of *multiple linear regression*. A number of such theories have been proposed but the emphasis here shall be on the Lens model (Brunswik, 1952), and Information Integration theory (N.H. Anderson, 1981).

Brunswik (1937, 1947, 1952) developed the Lens model to explain how perceivers use sensory cues to make inferences about the visual world. The model was adopted by Hammond (1955) and applied to the area of decision-making. The modelling techniques remain the same but the field is commonly known as Social Judgement Theory (SJT) when applied to decision-making and judgement-revision. Applications have been made to business, including policy making (Brehmer & Brehmer, 1988) and recruitment of salesmen (Roose & Doherty, 1976), psychology, including the performance of clinical psychologists (Hammond, 1955) and learning and inference (Klayman, 1988), medical judgements, including judging malignancy of ulcers (Slovic, Rorer & Hoffman, 1971; see Wigton, 1996 for a review) and general decision-making (Phelps & Shanteau, 1978; Brehmer, 1990).

The Lens model proposes that for some perceptual target that needs to be identified or judgement that needs to be made, there will be a number of cues providing information concerning the target or judgement that are used to make one's final decision. To illustrate the model, using an example provided by Beach (1997), a company may wish to identify the characteristics of a good salesperson. The identified characteristics would act as the cues in making decisions about prospective candidates, with cues given greater prominence or importance weighted more heavily.

51

Finally, the data is analysed using the statistical technique of multiple regression. A consistent finding is that the model performs better than the human judge (Goldberg, 1970; Dawes, 1971, 1979; Dawes & Corrigan, 1974; Camerer, 1981).

There are some interesting features of SJT that are relevant to contemporary theorising in the thinking and reasoning field. Brunswik emphasised *functional probabilism*, that is, an organism cannot know the environment with certainty and therefore, the relation between the organism and the environment must be considered in probabilistic terms. Such a view was counter to the approach of the day which emphasised general, univocal laws. Today, however, considering human behaviour in probabilistic terms is the norm (Oaksford & Chater, 1994; Evans & Over, 2004). An adaptation of SJT has been Cognitive Continuum Theory (CCT) (Hammond, 1986, 1996) which addresses the question as to why some tasks appear easy while others are difficult. A continuum is proposed from intuitive to analytical cognition. A task such as syllogistic reasoning involves the more demanding analytical cognition while a less cognitive demanding task, such as selecting candidates for graduate admissions, would involve intuitive cognition. The analytical-intuitive distinction has a number of parallels in the field of psychological reasoning (e.g. Evans, 1989; Evans & Over, 1996b).

Social judgement theorists, such as Doherty & Kurz (1996), admit that when making ratings participants may reduce the actual number of cues given to a cognitively manageable number. Thus, analytical cognition may be replaced by an intuitive mode of thought. This may also apply in situations involving logical or statistical problems

that are unfamiliar but whose content is familiar. Hence, SJT theorists emphasise the importance of task parameters.

The Lens model and SJT take a positive view of human cognition, compared to research on logical reasoning and heuristics and biases. A further advantage of the approach is that multiple regression analysis is robust. However, the Lens model in particular has been criticised for being too simplistic and the appropriateness of multiple regression has been questioned. Multiple regression analysis presumes that cues are additive with a candidate accumulating an ever higher score as more positive cues are added. Further, cues are treated independently with multiple regression, when cues may in fact interact, that is be multiplicative. Modifications can be made to allow for multiplicative combination, however, the procedure is cumbersome and seldom works (Kort, 1968). Multiple regression also assumes a linear relationship and has difficulty with relationships that are curvilinear. The final disadvantage is that SJT techniques presume a single level of inference, from the cues to the decision, ignoring potential intermediate levels of the decision process (Beach, 1997).

An alternative approach also employing multiple regression is *Information Integration* (N.H. Anderson, 1970). The decision maker is believed to integrate cued information either additively, multiplicatively or by averaging. An advantage of Anderson's theory is that the cues need not be quantitative, they can be qualitative. When averaging, the decision maker gives each cue a value, adds the values of each cue and divides by the number of cues. N.H. Anderson (1981) successfully applied averaging to studies of impression formation. Additivity is observed when cues are added producing a linear relationship when plotted in a graph. Alternatively,

multiplicativity is observed in a *fan* representation when cues are multiplied. As with other multiple regression approaches, weights may be applied to cues which are then multiplied by the assigned values. The integration of information can be applied to situations requiring judgement revision and such an approach is taken in the research reported in this thesis.

The theories of decision-making discussed above included a theory, SEU theory, that appears to be beyond the cognitive capabilities of humans. The Lens model offers a more viable explanation of human decision-making with its emphasis on environmental factors and also has the advantage of having similar features to theories of human reasoning. Information integration theory is also a viable contender, particularly when explaining judgement revision.

2.7 Deontic Reasoning and Decision-Making

Amongst the earliest theorists to propose that reasoning and decision-making were invariably linked was Evans, Over & Manktelow (1993). There may be distinctive differences between the two traditions which have resulted in separate fields of study, however, both reasoning and decision-making are components of rationality. As they state, "reasoning tasks usually involve making decisions" and "[i]n real-world situations the distinction between reasoning and decision-making is blurred" (Evans et al., 1993, p.166). There may be at least two forms of rationality, as suggested by Evans & Over (1996b); *Rationality 1* and *Rationality 2*. The former allows one to achieve one's goals, while the latter conforms to normative reasoning and decision-making systems. Empirical studies of reasoning (e.g. syllogistic reasoning) and decision-making (e.g. gambling behaviour) have usually tested Rationality 2, whereas

experimental studies of deontic or practical reasoning, and everyday reasoning test Rationality 1, that is, they involve goal attainment. Practical reasoning, involving the inferring of actions that ought or may be taken in a given situation, is a particularly pertinent example of the association between reasoning and decision-making. Further similarities between the two fields are the pragmatic factors of probability, utility, and social perspective, traditionally areas of concern for decision-making but which have recently been found to influence deontic reasoning, as discussed above. Not surprisingly, Evans et al. reject normative accounts of decision-making, not least for their inadequacies at explaining personal preferences during practical decision-making, preferring the notions of *bounded rationality* (Simon, 1983) and *satisficing* (Newell & Simon, 1972) in which the decision-maker finds an adequate solution, but not necessarily the optimal solution, within given resource constraints. Such a view is adopted in the thesis.

Other theorists have also turned to theoretical aspects of decision-making for explanation. Oaksford & Chater (1994) now consider deontic versions of the Wason selection task as involving decision-making as utility must be incorporated into their Bayesian approach if their theory of *information gain* can adequately explain such tasks.

The above review demonstrates there is much overlap between the research interests of those working in the field of reasoning and those in decision-making. Cross-fertilisation of ideas, methods and theorizing has begun, however, there is much more potential and this thesis makes an attempt to develop on these modest beginnings.

2.8 Rationale to the Thesis

The review of the literature in this and the previous chapter has shown the inadequacies of the normative approach to reasoning and decision-making. Such an approach assumes a logical capability beyond that observed in experimental research. Studies with syllogisms have shown many inconsistencies and biases in the reasoning of individuals with some syllogisms being solved readily in some contexts but not others. Syllogistic reasoning has been shown to be influenced by the subtleties of content, context and the understanding of the terms, All, Some, Some…not, and No.

The abstract selection task is a simple yet effective means of demonstrating that the truth table for material implication is not held in the heads of individuals. The thematic versions of the selection task, however, have acted as a bridge between the normative approach and the more realistic endeavour of investigating everyday reasoning and decision-making. However, a realistic context per se does not guarantee *logical* reasoning on the selection task (Manktelow & Evans, 1979). Further, a complete explanation of the Wason selection task has eluded researchers for decades, which has led some to question whether it is a suitable means for the investigation of human reasoning and rationality. Thus, syllogistic reasoning and deductive reasoning on the Wason selection task are prone to biases and were, therefore, rejected as methods of studying reasoning and decision-making in this thesis. The distinction between deduction and induction is increasingly considered as fallacious (Manktelow, 1999) and this afforded the opportunity to use a series of novel inferential reasoning and decision-making tasks throughout the experiments reported in this thesis. Evans (2002) has gone so far as to propose that the deductive paradigm is no longer the appropriate approach for studying the processes of reasoning and rationality.

Deontic reasoning as a discrete form of cognitive activity has been established for some time (Cheng & Holyoak, 1985) and several intervening factors have been found to influence the conclusions drawn when undertaking such reasoning, in particular the *perspective effect* (Gigerenzer & Hug, 1992; Manktelow & Over, 1991; Politzer & Nguyen-Xuan, 1992) and probability (Kirby, 1994; Manktelow et al., 1995; Stevenson & Over, 1995). Facilitation on thematic versions of the selection task, that is the selection of logically correct card choices, has also been demonstrated with these manipulations (Griggs & Cox, 1982; Johnson-Laird, Legrenzi & Legrenzi, 1972). This thesis explores further pragmatic influences of deontic reasoning, including scale of violation, mitigating and aggravating circumstances and power of source with the use of inferential reasoning and decision-making tasks.

Several theories of deontic reasoning have now been proposed and established as leading contenders for explaining this important phenomenon, including Pragmatic Reasoning Schema theory (Cheng & Holyoak, 1985), Social Contract theory (Cosmides, 1989), Mental Models theory (Johnson-Laird & Byrne, 1991; Manktelow & Over, 1995), Dual Process Theory (Evans & Over, 2004; Stanovich, 1999), Information Gain theory (Oaksford & Chater, 1994) and Probability theory (Over, Maktelow & Hadjichristidis, 2004). The investigation of pragmatic manipulations using inferential reasoning tasks will allow for the testing of these theories. The approach taken in this thesis assumes that reasoning and decision-making are associated and uses a simple procedure of sequential presentation to convert an inferential reasoning task to a decision-making task. Evans, Over & Manktelow (1993) were amongst the first to suggest the relationship between reasoning and decision-making using a dual factor theory. They proposed that much of everyday

reasoning and decision-making involves what they called Rationality 1, reasoning associated with one's goals, rather than the more systematic and logical Rationality 2. With the development of dual process theory the terms, System 1 and System 2 are now preferred. The pragmatic factors investigated in this thesis are considered as direct manipulations of System 1 in which inferences will be drawn from personal experience and knowledge with the given, familiar context.

The research reported in this thesis explores pragmatic factors of deontic reasoning using scenarios that would be familiar to most adults, such as employment rules and motoring offences. Rules were presented as if…then conditional statements, with the scenario elaborated to include rule violations. Participants made judgements of the seriousness of the violation on a verbal or numerical rating scale. Violations differed in type (e.g. speeding versus drink-driving), in scale (e.g. minor versus major) and in surrounding circumstances (e.g. mitigating versus aggravating). Judgement revision was investigated within the context of motoring violations by presenting details of the scenario serially. Violations are frequently used in deontic reasoning tasks, however, manipulations of scale have not been explicitly investigated before.

The authority or the power of the source of a statement was also studied in both formal (employment and armed forces) and informal (family and health) settings. The power of source experiments were developed by embedding conditional statements within promises, threats and warnings and by the use of conjunctive and disjunctive statements. Finally, serial presentation of the power of source scenarios allowed for judgement revision to be investigated with inducements (i.e. promises, threats and warnings). This research is novel in a number of respects. The authority making a

statement or inducement has not been directly manipulated in previous studies. The paraphrasing of conditional statements to the conjunctive or disjunctive form investigates if similar judgement ratings are made irrespective of the type of statement; an approach not taken before.

In sum, the rationale of the thesis was to investigate reasoning and decision-making in deontic contexts using inferential tasks. A simple manipulation from end-of-series to serial presentation allowed a reasoning task to be converted to one of judgement revision. An underlying assumption of the research was that reasoning and decision-making are invariably linked, involving similar cognitive processes and susceptibility to the same pragmatic manipulations. The factors of scale of violation, circumstances and power were investigated using familiar social contexts in direct manipulation of System 1 thinking.

Chapter 3

Scale of Violation in a *Working Rule* Scenario as a Pragmatic Factor of Deontic Reasoning

3.1 Experiment 1: Rationale

In this chapter it is proposed that *scale of violation* acts as a pragmatic factor within a deontic reasoning context. Two experiments are described in which scaled violations for a working hour rule are manipulated using deontic reasoning tasks. Two alternate methods of investigating scaled violations are the use of an inferential reasoning task and the large array selection task (LAST), as devised and implemented by Manktelow, Sutherland & Over (1995), and both are adopted in Experiment 1 and Experiment 2, respectively. We shall first consider the rationale for Experiment 1.

An adapted version of Gigerenzer & Hug's (1992) 'day off' rule provided the inferential reasoning task for Experiment 1. The 'day off' rule states, "If an employee works on the weekend, then that person gets a day off during the week", and was used by Gigerenzer & Hug within a selection task to demonstrate bilateral perspective effects. That is, different card selections are observed when participants are asked to take either the perspective of an employee or the employer. Those taking the perspective of the employee were cued to the potential violation of the 'day off' rule by the employer and were more likely to select the p and not-q cards (i.e. the cards stating, "Worked at the weekend" and "Did not get a day off in the week", respectively). Conversely, when cued to the perspective of the employer and the potential violations on the part of employees, participants were more likely to choose the not-p and q cards (i.e. the cards stating, "Did not work at the weekend" and "Took a day off in the week", respectively). The 'day off' rule was one of several scenarios

used by Gigerenzer & Hug to disentangle two theoretical aspects of Cosmides' (1989) Social Contract Theory: a rule as a social contract and the proposed cheater-detection algorithm.

In the current research the 'day off' rule is used to investigate the effects of scaled violations. Research studies of rule transgression, typically conducted using a selection task, tend to consider violations as a binary relation, that is, as a violation or non-violation. Influential theories of deontic reasoning, particularly schema theories, appear to implicitly hold this view as the detection of a violation is represented as an all-or-none effect. However, an inferential reasoning task has the advantage that it is relatively easy to provide participants with graded options unlike the standard selection task which offers only four choice options. Admittedly, the inferential task changes the nature of the problem to that of a judgment or decision-making one. An inferential task can be said to involve an evaluation prior to making a judgment or decision whereas the selection task requires one to identify an exception to a rule; a violator in the context of deontic scenarios. Furthermore, an inferential reasoning task does not accord with a specific correct response, nor is it linked to a normative standard or theory, such as logic. Performance on the selection task has been associated with a number of factors, including individual differences, for instance if the individual has received training in formal logic and/or has a high I.Q. (Stanovich & West, 2000) or has training in mathematics (Gigerenzer & Hug, 1992). Thus, an alternative, non-traditional method was undertaken to investigate individuals' reasoning with scaled violations within deontic contexts.

Two hypotheses were predicted for this experiment:

1. The allocation of fines by participants would significantly increase as the scale of violation increased over Day 1 (1, 4, & 8 hours) for a *working rule* task.

2. The allocation of fines by participants would significantly increase as the scale of violation increased over Day 2 (1, 4, & 8 hours) for a *working rule* task.

3.2 Experiment 1: Method

Participants

Participants comprised of 13 part-time students attending an evening 'A' level Psychology class at Stafford Further Education College. Ten participants were female and 3 male with an age range of 16-47 years (median age 23 years). Two female participants declined to give their age. All participants were volunteers and had received no tuition in the area of thinking and reasoning. No further demographic details were obtained.

Materials

The materials comprised of an instruction sheet and eight individual employee work cards. The instruction sheet, given to each participant, was presented on A4 paper and provided a set of instructions for the task along with an example.

The set of employee work cards were of A5 size with participants instructed to inspect each card individually. The employee cards consisted of 2 controls in which no violation of the rule had occurred, 3 cards showing a violation of the rule for Day 1 (incremented for 1,4, & 8 hours over-worked), and 3 cards for violations of Day 2 (incremented for 1, 4, & 8 hours worked on what should have been a day off). For a copy of the instruction sheet and materials see Appendix A1 (Experiment 1:

Materials).

Design

A within-participants design was implemented. Two independent variables (IV's) were manipulated, scale of violation and day on which violation occurred. Scale of violation comprised of three levels, 1 hour, 4 hours, and 8 hours over-worked. Day of violation comprised of two levels, Day 1 and Day 2. A control condition was also employed in which no violation had occurred. Participants completed all conditions of the experiment. The dependent variable was the fine allocated to an employee by the participant.

Procedure

The experiment was undertaken in one sitting with participants presented with an instruction sheet and a set of 8 employee cards. Participants were asked to read the instructions carefully, follow the given example and to ask any questions, if necessary.

The instructions primed the participant to imagine that they were a Personnel Manager working for a local ambulance service. The ambulance service was said to be concerned that some employees were working over their scheduled hours, therefore, a new rule had been introduced and it was their task as Personnel Manager to investigate if the new rule was being adhered to. The rule stated, *If drivers work for more than 8 hours in a day, then they must take the next day off*. An example Employee Work Record card was shown in which the rule had been violated; the employee had worked over their scheduled hours on Wednesday (Day 1) and should

have taken the following day off. Participants were instructed that employees must be fined if the rule is violated and, as the Personnel Manager, it is their responsibility to determine the fine from a given range. The fines were incremented as follows:

£0 (No fine), £20, £40, £80, £160, £320 (One weeks' wages)

The *working rule* was applied to eight employees, whose details were presented on individual employee cards. Two cards acted as the control condition in which no violation of the working rule had occurred. One of the control cards showed the employee having worked their scheduled hours (i.e. 8 hours per day) and the other showed an employee working over-hours on one day and taking the following day off, thus complying with the given rule. The control condition also indicated if the task had been understood by the participants as no fine should be applied.

Two experimental conditions were employed involving violations of the *working rule*. A violation required that an employee must work for more than 8 hours on Day 1 and to work for any number of hours on Day 2. Thus, the conjunction of working over 8 hours on Day 1 and working for at least 1 hour on Day 2 must be observed for a violation to have occurred; if only one of these conditions were met then no violation was said to have occurred.

The first experimental condition comprised of 3 employee cards representing potential violations of the rule on Day 1. These cards were incremented for 1, 4 & 8 hours over-worked and showed the employee having worked 9, 12, & 16 hours, respectively. The following day the employee should have taken the day off but the

work record clearly stated that they had worked an 8-hour day. The number of hours over-worked were also shown on the employee work record and an asterisk indicated that an employee had worked on what should have been a day off, and thus a violation of the rule had occurred. The purpose of the asterisk was to highlight the occurrence of a violation and prevent the task from being too cognitively demanding for the participant. Henceforth, this experimental condition is labeled, *Day 1*, reflecting the scaled number of hours worked above the contractual 8 hours and the potential violation if the employee were to work on the subsequent day.

A second experimental condition comprised of 3 employee cards representing violations for Day 2, incremented for 1, 4 & 8 hours worked on what should have been a day off. Each of these cards showed the employee had worked for 10 hours on Day 1 and should, therefore, have taken the following day off. However, the working rule was violated because the three employees worked 1, 4 & 8 hours, respectively on what should have been a day off. The number of hours over-worked (2 hours) were clearly shown on the employee work card and an asterisk indicated the employee was working on what should have been a day off. This experimental condition is henceforth labeled, *Day* 2, representing the scaled number of hours worked on Day 2 and the fact that a violation has occurred due to the employee having worked over 8 hours on the previous day.

Participants were asked to indicate the rate of fine on each of the eight individual employee work records, including no fine (£0). No time limit was set for the task.

3.3 Results

For raw data see Appendix B1 (Experiment 1: Raw Data). The control condition, in which no violation occurred, produced fines on only two occasions and was, therefore, excluded from the analysis. One participant attributed £0 (no fine) to all employee work records, however, this was deemed acceptable and the data was retained for analysis.

Table 3.1 shows the mean fines and standard deviations for the two experimental factors; scale of violation and day of violation. It can be observed that there is an increase in mean fine as scale of violation increases for both Day 1 and Day 2, with the effect being much greater for Day 1. There is some variation in the dispersion of fines as shown in the standard deviation (S.D.). This is particularly the case for the Day 1 condition where the S.D. for 8 hours worked over is extremely high, reflecting the spread in participant fines from £0 - £320.

	Day 1 hours worked over			Day 2 hours worked over		
	1 hour	4 hours	8 hours	1 hour	4 hours	8 hours
Mean	17	42	92	18	23	28
S.D.	11.1	30	89.3	15.2	22.9	22.4

Table 3.1: Mean fine and standard deviation (S.D.) for violations of working rule over 1, 4 & 8 hours worked over on Day 1 and Day 2

Statistical analysis was undertaken on the data using SPSS (version 15). A two-way (2x3) within-subjects analysis of variance (ANOVA) was used to test for statistical

significance. The principle of homogeneity of variance was shown to be violated due to the differences in the S.D. scores. Further, Mauchly's Test of Sphericity was significant for the scale of violation factor and for the interaction, scale x day. Therefore, the Greenhouse-Geisser Epsilon correction for degrees of freedom was used for reporting purposes.

As predicted a significant difference was found for scale of violation ($F(1,24) = 11.086$, $p = 0.004$). A significant, but unpredicted, effect was observed for the day of violation (Day 1 v. Day 2) ($F(1,12) = 11.327$, $p = 0.006$) and a significant interaction effect (scale x day) was also found ($F(1,24) = 6.156$, $p = 0.025$).

The interaction effect is illustrated in Fig. 3.1 (below). At 1 hour violation similar rates of fines are observed for both days, but whereas Day 2 shows a negligible increase over 4 and 8 hours, Day 1 produces sharp rises for 4 hours and 8 hours worked over. Hence, participants appear to view the potential violation on Day 1 as more serious than the subsequent hours worked over on Day 2 when an actual violation has occurred.
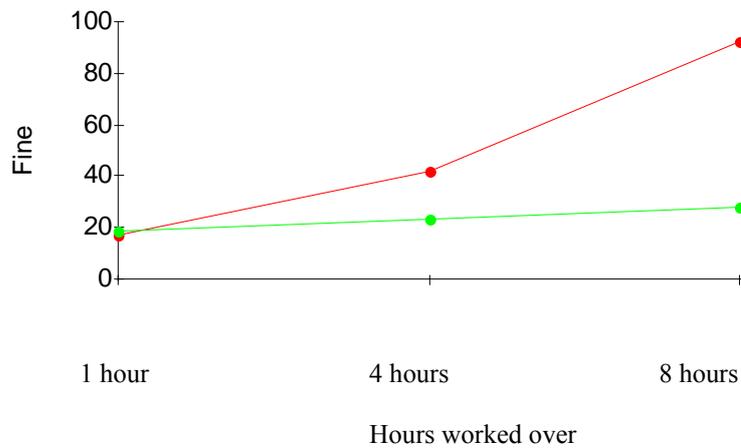
Figure 3.1: Interaction effect showing mean fines allocated to employees for working 1, 4 & 8 hours over for Day 1 (red) and Day 2 (green)

3.4 Experiment 2: Rationale

Experiment 1 established scale of violation as an influential pragmatic factor within a deontic context when a task is embedded in an inferential reasoning task. The aim of Experiment 2 is to investigate whether the findings can be transferred to a deductive reasoning task.

Scale of violation requires a series of graded options to be presented to participants and the large array selection task (LAST) allows this. Manktelow et al. (1995) used this technique with an adapted version of Cheng & Holyoak's (1985) 'immigration rule'. The rule stated, "If a person has ENTERING on one side of their immigration form, then they must have CHOLERA on the reverse side". Participants were cued to the role of an immigration officer and asked to detect violators of the rule. A probabilistic variable was incorporated by including information that cholera was

more prevalent in the tropics and by including the country of origin of each immigrant. Note that Manktelow et al.'s participants were specifically primed to the increased probability of cholera being brought into the country; no such prime was given in this experiment.

The LAST, therefore, affords the opportunity to consider whether the findings from an inferential task are transferable to a deductive reasoning task and to compare any performance differences between the two forms of task.

Three hypotheses were predicted for this experiment:

1. Card selections will significantly differ between the Control condition and potential violation on Day 1.

2. Card selections will significantly differ between the Control condition and potential violation on Day 2.

3. Card selections will significantly differ between the potential violation on Day 1 and the violation on Day 2.

The third hypothesis was proposed due to the finding in Experiment 1 of participant's perceived difference between the potential and actual violation of Day 1 and Day 2, respectively.

3.5 Experiment 2: Method

Participants

Forty-two full-time students attending Stafford Further Education College acted as participants. The participants comprised of 10 males and 32 females with an age range of 16-28 years (median age 18 years). Participants were randomly allocated to one of

three conditions; control, violation over Day 1, and violation over Day 2. One participant from each condition declined to give their age. The control condition comprised of 12 participants (4 male and 8 female) with an age range of 17-18 years (median age 18 years). The Day 1 experimental condition consisted of 19 participants (6 male and 13 female) with an age range of 16-20 years (median age 17 years). The Day 2 experimental condition comprised of 11 female participants with an age range of 17-28 years (median age 18 years).

Materials

The materials consisted of an instruction sheet and three large array selection tasks (LAST's). The three large arrays included; a control array comprising non-incremental cards; an experimental array consisting of graded p cards i.e. hours worked over on Day 1; an experimental array showing graded not-q cards i.e. hours worked over on Day 2. To view a copy of the instruction sheet and materials see Appendix A2 (Experiment 2: Materials).

The instruction sheet primed the participant to the role of Personnel Manager with a local ambulance service, as in the previous experiment. The rule to be investigated stated, *If drivers work for more than 8 hours on a Friday, then they must take Saturday as a day off*. An example Employee Work Record was given, showing an employee having worked for 9 hours on a Friday and having taken the following day, Saturday, as a day off. The participant was instructed that their task was to inspect a set of Employee Work Records and to indicate, by placing a tick in the box placed below the employee record, if they wished to turn the employee card. The placing of a tick indicated a potential violation of the working rule.

The control group received an array consisting of 12 none graded employee work record cards. Three cards (the p cards) showed an employee having worked on Friday for 10 hours; three cards showed employees working Friday for 7 hours (the ¬p cards); a set of three cards showed employees taking Saturday as a day off (the q cards); and three cards represented employees working for 8 hours on Saturday (the ¬q cards).

The experimental condition 1 group had an array in which the hours worked varied over Day 1 (Friday). A set of 12 employee work records were displayed with the p cards showing three employees having worked 10, 12 & 16 hours on day 1 (Friday) and the ¬p cards stating three employees had worked 4, 5 & 7 hours on Friday. The three q cards displayed the employees as having taken Day 2 (Saturday) as a day off, and the ¬q cards showed three employees working for 8 hours.

The array for experimental condition 2 comprised of 12 employee work records with graded hours worked on Day 2 (Saturday). The p cards showed the employees having worked for 10 hours on Day 1 (Friday) and the ¬p cards the employees working for 7 hours on Day 1. The three q cards stated the employees had taken Day 2 (Saturday) as a day off and the ¬q cards showed three employees having worked for 2, 5, & 8 hours on Day 2.

A summary of the cards presented in each array of the three conditions is given below:

Control Condition (none graded p or ¬q cards)

p cards	3 x 10 hours*
not-p cards	3 x 7 hours
q cards	3 x DAY OFF
not-q cards	3 x 8 hours*

Experimental Condition 1 (graded p cards)

p cards	10, 12 & 16 hours *
not-p cards	4, 5 & 7 hours
q cards	3 x DAY OFF
not-q cards	3 x 8 hours *

Experimental Condition 2 (graded ¬q cards)

p cards	3 x 10 hours *
not-p cards	3 x 7 hours
q cards	3 x DAY OFF
not-q cards	2, 5 & 8 hours *

NOTE: * denotes potential violation of rule and, therefore, correct card selections;

Design

A between-participants design was employed with each of three conditions tested on a separate group of participants. A control group judged a set of non-incremental

employee work cards while two experimental groups of participants considered potentially violating employee cards. Experimental group 1 judged a set of incremental p cards representing three levels of a potential violation on Day 1 (10, 12 & 16 hours worked on Day 1). Experimental group 2 judged a set of graded q cards representing three levels of potential violation (2, 5 & 8 hours worked on a Day 2). Hence, two independent variables (I.V.'s) were employed; graded p cards and graded q cards, representing potential violations for Day 1 and Day 2, respectively. The dependent variable was the selection of a card(s) by the participant, indicating a potential violation of the working rule.

Procedure

Participants were tested in groups with each of the three conditions tested separately. Each participant received an instruction sheet and an array of 12 employee work record cards, both presented on A4 paper. Participants were asked to read the instructions and time was made available for questions and clarification, if required. As Personnel Manager it was the participant's task to inspect the array of Employee Work Record cards to ensure that a given working rule had not been transgressed. Only one side of each employee work record was visible and, therefore, the participant's task was to select those cards were they believed a potential breach of the rule may have occurred.

The normative solution to the selection task and the LAST is the selection of the p and ¬q cards and this is what was deemed correct in this instance, as these represented potential violations of the working rule. However, all card selections for each participant were recorded to allow a comparison across the three separate conditions.

3.6 Results

For raw data illustrating the individual card selections of each participant see
Appendix B2 (Experiment 2: Raw Data).

Table 3.2 shows the percentage of card selections for each of the three conditions of
Experiment 2. The percentages were derived by adding the number of card selections
made by participants and dividing by the possible total number of cards. For example,
for the control condition a total of 33 p-cards were selected by participants from a
possible overall total of 36 (i.e. 12 participants x 3 cards). With unequal participant
numbers across conditions, frequencies of card selections were not a feasible form of
analysis.

----

	p card*	¬p card	q card	¬q card*
Control	92	6	58	58
Day 1	88	0	44	56
Day 2	94	0	27	61

**Table 3.2: Percentage card selections for each of the three conditions of
Experiment 2; * denotes correct card choices**

The p card was the most selected card for all three conditions and the not-p card the
least selected card. This finding is consistent with the majority of selection task
observations. The not-q card, also considered a correct choice, was selected to a lesser
degree but equally chosen was the q card. The incremental p cards (Day 1) show
similar card selections to the baseline group but with the q card selected slightly less.

Day 2, which contained incremental not-q cards, produced the highest percentage rates of card selections for the p and not-q cards. However, the q card was selected less frequently.

A series of complex Chi-square tests were conducted to compare card selections between the three conditions. Calculations were made using the Chi-square Calculator software (see Appendix D2 Experiment 2: Chi-square output). For Control x Day 1 x Day 2 a significant difference in card selections was found ($X^2 = 20.74$, df = 6, p < 0.01). Further analysis revealed a significant difference between the Control and Day 2 conditions ($X^2 = 14.91$, df = 3, p < 0.01). However, non-significant findings were observed for Control x Day 1 ($X^2 = 6.39$, df = 3, p = 0.10) and for Day 1 x Day 2 ($X^2 = 4.39$, df = 2, p = 0.20).

These findings support the hypothesis that card selections will significantly differ between the Control condition and scale of violation on Day 2. However, the other hypotheses comparing the Control condition with scale of violation on Day 1 is not supported. Further, the observation of a perceived difference between potential violations on Day 1 and Day 2 in Experiment 1 was not repeated in this study.

3.7 Discussion

The results from Experiment 1 appear to suggest that participants' judgments of violation are primarily influenced by the initial rule-breaking and by the scale of violation, but not by the scale of the subsequent violation. That is, working over the specified hours on Day 1 is judged an offence and the greater the violation the more serious the offence. However, the scale to which the rule is abused on the following

day (Day 2) is not judged as such a serious infringement. This was an unexpected finding and one that has not been reported in the literature before, indicating that human reasoning with violations is more complex than hitherto realised.

The interaction effect shown in Figure 3.1 is of particular interest because it clearly shows the seriousness of the working rule transgression is judged to be the hours worked over on Day 1; the number of hours worked on Day 2, a day of rest, has a negligible effect on the perceived violation. This is surprising because a violation of the rule requires both overworked hours on Day 1 and hours worked on Day 2 and one might expect greater sensitivity to the numbers of hours worked on what should have been a day off (i.e. Day 2). A possible explanation for the increased fine observed in the Day 1 condition could be a *primacy effect* as discussed below. Alternatively, the understanding and interpretation of rules and their transgression may be a factor. This would suggest the cognitive processing of rules and violations is important, a view considered by Dominowski (1995) in relation to findings with the Wason selection task. A more complete understanding of this unpredicted finding would require a systematic investigation of the processes involved in the comprehension of everyday rules.

For Experiment 2 only one significant result was obtained; scale of violation does not appear to influence card selections on a large array selection task. As the *correct* cards (p & not-q) tended to be selected by the majority of participants across conditions, it could be that the task is just too simple or else it omits an important element in studying judgmental behaviour. Alternatively, participants may be ignoring the subsequent violation i.e. the q and not-q cards because such a violation seems

unlikely, especially in the case of the employee working for 16 hours on day one, where it may seem highly improbable that the employee would arrive for work the following day. Such a view has some support from the finding that the p card is selected more frequently than both the q and not-q cards across all three conditions. If this suggestion has merit it could be compared to Kirby's (1994) finding that *age* determines the probability of an individual being *checked* for drinking alcohol and could suggest *scale of violation* is also a probabilistic factor in reasoning.

Studies of deontic reasoning have relied on a number of manipulations and the experiments conducted here are no exception. A ubiquitous factor in thematic versions of the selection task that consistently elicits logical card selections is *role playing*. A wide range of roles have been used including, an anthropologist (Cosmides, 1989), mother and son (Manktelow & Over,1991), customs officer (Manktelow et al.,1995) and doorman (Kirby,1994). Perspective is a further factor that has been found to facilitate card selections, for example, Manktelow & Over (1991) used a mother and son scenario, and Gigerenzer & Hug (1992) used a variety of bilateral perspectives, including employee-employer within the 'day off' rule. Even deontic tasks with children have tended to include perspective and role-playing in the experimental situation (eg. Cummins, 1996a, 1996b; Girotto, Light & Colbourn, 1988; Light, Girotto & Legrenzi, 1990).

Participants in these experiments were cued to the role of a personnel manager but the bilateral perspectives of employee - employer were not investigated. Manktelow & Over's (1991) notion of *utility* can be applied to the findings obtained in Experiment 1. From the perspective of the personnel manager it is to their benefit to identify

violators; indeed in this instance it is one of their key roles. What is not made explicit in the working hour scenario is the benefit to the employee of working over hours on what should be a day off. From ones knowledge of working practices participants may have assumed that working over hours would result in the payment of overtime or that it was a genuine act of good will. Therefore, although one can ascribe benefits to each party, bilateral utilities are not made fully transparent in this particular context.

Comparing Experiment 2 and Manktelow et al.'s LAST study there is a further noticeable difference in the instruction given to participants. Manktelow et al. cued participants to the high probability of cholera as a tropical disease. No such cue was used in this research which may contribute to the different findings with the LAST. There is also Evans' (2002) view that the *deductive paradigm* may not be appropriate when dealing with particularly realistic content, as is the case here.

As stated, Experiment 1 produces an unexpected finding in that scale of violation over Day 1 produced significantly greater fines than comparable violations over Day 2. There is no known theoretical reason for such a finding. One may speculate that what is being observed is a *primacy effect*, which has been investigated extensively in several areas of psychological literature. Within the field of *impression formation*, arguably the most researched in relation to this effect, three prime theories have been proposed to explain the prevalence of the primacy effect. N.H. Anderson (1974) has suggested *attention decrement*, that is, individuals attend to the initial information but fail to attend to subsequent information. This may apply here as participants are cued by the rule to the potential violation on Day 1. Alternative theories of the primacy effect include those of Asch (1946) who proposed that subsequent information is

made consistent with the primacy information and the suggestion of Luchins (1957) that recency information is discounted or dismissed. These latter two theories require participants to attend to the most recent material and we may expect discrimination in the levying of fines if attention was given to the violations over Day 1 and Day 2. Consideration of the findings in terms of a primacy effect, suggest that the hitherto unassociated areas of human reasoning and social cognition are not as mutually exclusive as the research literature would imply.

In summary, *scale of violation* may be considered a further probabilistic variable, similar to the way in which Kirby (1994) uses age in the drinking-age rule and Manktelow et al. (1995) use country of origin within the immigration rule. However, a contrasting finding between the two experiments reported here was that the effect of scale of violation was only observed in the inferential task, and not when embedded within a LAST. A possible explanation for the non-significant findings with the LAST, a deductive reasoning task, is that participants may quite easily identify violators. However, to discern how the individual discriminates and reasons with scaled violations in more realistic, deontic contexts one must employ an inferential reasoning task.

Chapter 4

Scale of Violation and Mitigating and Aggravating Circumstances as Pragmatic Factors of Deontic Reasoning

4.1 Experiment 3: Rationale

The previous chapter found scale of violation to influence judgments made with a working hour rule when embedded within an inferential reasoning task. It seems reasonable to propose that scale of violation is a general pragmatic factor, applicable to a range of deontic contexts, and having varying effects on ones reasoning and decision-making. Amongst the best examples of deontic rules are transgressions of the law, such as motoring offences.

Studies of deontic reasoning manipulate violations in various ways in order to test statements of permission and obligation. Most often a deontic task involves checking for violators (Cheng & Holyoak, 1985; Cosmides, 1989; Manktelow & Over, 1991). Other manipulations are possible, for instance, Kirby (1994) used scaled violation within a drinking rule context, however, this study involved only the potential of violation; for example, that a 19 year old may be drinking alcohol when the rule specifies that a person must be over 21 years old. Surprisingly, direct manipulations of scale of violation have not been reported in the literature. The experiments reported in this chapter use realistic transgressions, motoring offences, in which scale of violation is explicitly manipulated.

In addition to the transgression itself, is the circumstances surrounding the violation, in particular mitigating and aggravating circumstances which are permissible as evidence in the court of law. It is expected that some circumstances will mitigate the

seriousness of a transgression, while others will increase the seriousness and associated penalties. There is evidence from the deductive reasoning literature showing how logically drawn conclusions can be suppressed when additional premises are provided (Byrne, 1989). Byrne used a familiar, everyday context (studying in a library) and one would expect that similar findings would be achieved with the use of violations presented within an inferential reasoning task.

A number of experiments have demonstrated the non-monotonicity and defeasibility of conditional reasoning when alternative causes and disabling conditions (Cummins, 1995; Cummins, Lubart, Alksnis, & Rist, 1991; Fairley & Manktelow, 1997; Fairley, Manktelow, & Over, 1999) are provided. According to truth-functional logic, if the premises of a conditional argument are true then the entailed conclusion must logically follow, that is, conditional reasoning is non-monotonic. Also, further information in the form of an additional premise cannot alter the valid conclusion, that is, the valid conclusion of a conditional is not defeasible. However, there is strong empirical evidence for the defeasibility of logically valid conclusions as shown by the work of Byrne (1989) and Stevenson & Over (1995).

Cummins (1995) and Cummins et al. (1991) have investigated cases that influence deductions made in the process of causal reasoning, which they call, *alternative causes* and *disabling conditions*. Alternative causes are defined as, "a cause (other than the one cited in the causal rule under consideration) that is capable of evoking the effect cited in the rule" (Cummins, 1995, p.647). Similarly, a disabling condition is, "an event that could prevent an effect from occurring in the presence of a viable cause" (Cummins, 1995, p.647).

81

Using causal rules presented in conditional, if …then, form, Cummins and colleagues set their participants two tasks: a generation task and a reasoning task. The generation task required the generation of alternative causes and disabling conditions for given causal conditional statements. For example, Cummins et al. (1991) used the following statements and asked participants to generate possible alternative causes:

Rule: If Joyce eats candy often, then she will have cavities

Fact: Joyce has cavities, but she does not eat candy often

Write down as many circumstances as you can that could make this situation possible

A second group were given a similar rule and asked to generate disabling conditions:

Rule: If Joyce eats candy often, then she will have cavities

Fact: Joyce eats candy often, but she does not have cavities

Write down as many circumstances as you can that could make this situation possible

It was predicted that some contexts involving causal conditional rules would afford a greater generation of alternative causes and disabling conditions than others. For instance, the examples given above should allow participants to generate alternative causes contributing to Joyce having cavities without eating candy, and reasons (disabling conditions) why Joyce may eat candy often but not have cavities. However, it may not be so easy to generate alternative causes and disabling conditions with other causal contexts. For instance, Cummins et al. (1991) cite the example, "If my finger is cut, then it bleeds". In this case it would be difficult to generate disabling

conditions where one's finger does not bleed when it is cut or alternative causes where one's finger is bleeding without being cut.

The findings of the generation task supported the hypothesis that context does indeed influence the number of alternative causes and disabling conditions that may be devised for causal conditional arguments.

The reasoning task asked participants to consider a sample of causal arguments taken from the generation task and to determine an acceptance rating for the conclusion of the argument. The materials were selected on the basis of there having been *many* or *few* alternative causes and disabling conditions afforded to the conditional arguments during the generation task. Conditional arguments were then created using the four traditional forms of argument; Modus Ponens (MP), Modus Tollens (MT), Affirmation of the Consequent (AC), & Denial of the Antecedent (DA). As hypothesized, MP and MT arguments were more likely to be accepted as true if they were associated with few disabling conditions; that is, fewer counter examples could be generated to explain the causal rule. The fallacies (AC & DA) were only influenced by alternative causes; few alternative causes lead to greater acceptance of the given argument and many alternative causes increased the likelihood of rejection of the argument. Cummins (1995) was able to demonstrate the same effect with reversal causal arguments. Typically, causal arguments would be presented in the form of cause followed by effect: Cummins reversed this relationship (effect followed by cause) and observed a reversal of the results: MP & MT arguments were influenced by the number of alternative causes and AC & DA arguments were influenced by the number of disabling conditions.

The studies of Cummins and colleagues demonstrate that with causal conditional arguments participants generate varying numbers of alternative causes and disabling conditions depending on the given context and furthermore, the acceptance or rejection of logical conclusions is dependent on the number of such causes and conditions.

Fairley, Manktelow & Over (1999) developed Cummins' work in three key ways: alternative causes and disabling conditions were applied to deontic contexts, *additional requirements* were included as a further potential modifying factor of causal reasoning, and materials were presented in the form of a either a Wason selection task or a large array selection task (LAST). The conditional precaution, "If you wear rubber gloves, you may clear up spilt blood" has been used effectively in a number of experiments investigating deontic reasoning (Cheng & Holyoak, 1985; Manktelow & Over, 1990). Fairley et al. used a variation of this rule, "If you wear rubber gloves, you are protected from infection", and embedded it in scenarios concerning hospitals. Participants were primed to either an alternative cause (effective antiseptics now exist reducing the need for wearing rubber gloves) or an additional requirement (protective overalls are required in addition to rubber gloves) and shown four card choices, representing the p, not-p, q, & not-q options of the precautionary rule. The modal card choices corresponded to the prediction that alternative causes would result in the greater selection of the not-p and q cards, while additional requirements would increase the selection of the p and not-q cards. This finding was shown with causal reasoning, however, as pointed out by Fairley et al. the card selections follow those of the so-called perspective effect observed in deontic

reasoning (see Chapter 2, Section 2.3). Fairley et al.'s interpretation of the findings were that alternative causes question the necessity of wearing rubber gloves, while additional requirements, act in a similar way to disabling conditions, by questioning the sufficiency of wearing rubber gloves.

In a second experiment a LAST was used allowing for much stronger predictions. Not only would the alternative causes condition expect greater incidences of not-p and q card selections but also the card showing the alternative cause should be preferentially selected. Similarly with disabling conditions (Experiment 2 incorporated disabling conditions rather than additional requirements), the p and not-q cards were predicted to be the most frequently selected cards and the card naming the specific disabling condition was also predicted to be preferentially chosen. The predictions were supported using a novel conditional statement, "If you study hard, then you do well on tests" (the alternative cause stated, "You think that test can be unfair, because some students cheat when they take them"; the disabling condition stated, "You think that tests can be unfair, because some students feel terribly nervous when they take them").

Both Cummins et al. and Fairley et al. explain their findings in terms of sufficiency and necessity and although they disagree on the finer points of their respective theories (see Fairley & Manktelow's, 1997 comments and Cummins', 1997 reply) they are agreed that alternative causes affect the acceptance of an inference in terms of the necessity of a cause-effect relationship, whereas disabling conditions affect the inferences drawn in terms of the sufficiency of a causal relationship.

There is, therefore, strong evidence for the influence of alternative causes, disabling conditions, and additional requirements in deontic as well as causal contexts. Given these findings we might expect that aggravating and mitigating circumstances would also evoke an influence on conditional reasoning, particularly in the deontic domain. A general statement regarding the law may be stated in the following conditional argument, "If you commit a transgression of the law, then you are liable to a fine". Aggravating circumstances would be predicted to increase the probability of the consequent coming true if the antecedent is true. Mitigating circumstances, however, would be expected to lower the probability of the consequent coming true if the conditions of the antecedent were met. There is strong evidence for doubtful premises and conclusions to be considered in probabilistic terms (George, 1995, 1997; Stevenson & Over, 1995).

Two experiments were conducted using the context of motoring offences. The motoring offences of speeding and drink-driving were chosen because they are transgressions that would be familiar to most adults. The first experiment (Experiment 3) involved participants considering the seriousness of either the speeding or drink-driving offence when the circumstances, classified as mitigating, aggravating or neutral, were embedded. It was predicted that a significant difference would be observed in the ratings of seriousness between the type of circumstances for both the speeding and drink-driving violations. One may also expect the two offences, speeding and drink-driving, to be perceived quite differently. Therefore, a second hypothesis predicted a difference in the ratings between the two motoring transgressions. In the second experiment (Experiment 4) participants judged whether a motorist should, ought or must be fined given the transgression and the

accompanying circumstances.

The following hypotheses were predicted for Experiment 3:

1. A significant difference will be found between the circumstances (mitigating, neutral, and aggravating) surrounding the motor transgressions of speeding and drink-driving.

2. A significant difference will be found between ratings of seriousness for the motoring offences of speeding and drink-driving.

4.2 Experiment 3: Method

Participants

Twenty undergraduate students attending the University of Wolverhampton acted as participants. No details regards age, gender or demography were taken. Ten participants rated the *seriousness* of a speeding car driver while considering a set of circumstances. A further 10 participants rated the same set of circumstances but considered a car driver under the influence of alcohol.

Materials

The materials comprised of two A4 size booklets; one booklet related to a car driver driving above the legal speed limit and the other described a driver driving whilst over the legal limit of alcohol.

The first page of the booklet provided standardized instructions for the task. The instruction was given to judge the seriousness of a motoring offence (either speeding or drink-driving) while considering additional circumstances. *Seriousness* was rated

on a scale of 1-10; 1 representing "not at all serious" and 10 "very serious". An example of the motoring transgression was given along with an additional circumstance, that the driver was "late for an important sales meeting". To avoid influencing participant ratings and suggesting socially desirable responses, two potential choices were shown; a response of 2 (classified as "not serious") and 9 (classified as "very serious").

The following pages of the booklet consisted of 36 circumstances that had been classified, prior to experimentation, as mitigating, neutral, or aggravating. The task for the participant was to rate the seriousness of driving above the legal limit or above the legal limit of alcohol under each of the 36 circumstances. See Appendix A3 (Experiment 3: Materials) for copies of the two booklets used in this experiment.

Design

A mixed experimental design was employed in which offence was a between-participants factor and circumstances a within-participants factor. Two levels of the transgression factor were used; speeding and drink-driving. The circumstances factor had three levels; mitigating, neutral and aggravating. The type of offence and set of circumstances acted as independent variables (I.V.'s) and the participants rating of seriousness acted as the dependent variable (D.V.).

Procedure

The experiment was undertaken in one sitting with 10 participants rating the speeding offence and a further 10 participants completing the drink-driving transgression. After reading the standardised instructions on the front of the booklet participants were

given the opportunity to ask questions in order to clarify the task. Each of the 36 circumstances was rated by each participant with no time limit set for the task.

The circumstances were classified under three headings prior to experimentation. Mitigating circumstances are situations that are expected to reduce the seriousness of an offence; neutral circumstances are considered unrelated to the transgression and, therefore, should have little or no impact on the rating of seriousness; and aggravating circumstances are expected to exacerbate an offence, thus increasing the culpability of the offender. While devising the circumstances consideration was given to a number of associated factors, such as *driving conditions* (e.g. visibility, conditions of road), *time of day* (e.g. 2am or 5pm, peak hour), *situational factors* (e.g. celebration, birthday), *lateness* (for work or aeroplane), *characteristics associated with the driver* (e.g. experienced or *joy rider*), and *other circumstances* (e.g. previous convictions or no previous convictions). Twelve circumstances were classified as mitigating, 12 as neutral and 12 as aggravating, providing a total of 36 circumstances. The circumstances were presented in a random order within each of the two booklets to avoid participants identifying the nature of the experiment. See Appendix A3 (Experiment 3: Materials) to view the order of the circumstances as presented to participants in the booklets and Appendix B3 (Experiment 3: Raw Data) for a classification of the circumstances.

Participants rated seriousness on a scale of 1-10 where 1 was described as "not at all serious" and 10 "very serious". The rating acted as the score for that particular circumstance given the motoring offence.

4.3 Experiment 3: Results

For raw data see Appendix B3 (Experiment 3: Raw Data). The task was completed satisfactorily by all participants and, therefore, all data was used in the analysis.

Table 4.1 shows the mean total ratings of seriousness and standard deviations obtained for the two motoring offences and the three types of circumstance.

OFFENCE	CIRCUMSTANCE		
	Mitigating	Neutral	Aggravating
Speeding	5.84 (1.44)	7.08 (1.38)	8.98 (0.81)
Drink-driving	9.00 (0.78)	9.44 (0.74)	9.67 (0.60)

Table 4.1: Mean *seriousness* ratings and standard deviations (in parentheses) for two motoring offences (speeding & drink-driving) across three conditions of circumstance (mitigating, neutral & aggravating)

An increase in the mean total ratings for the speeding offence across mitigating, neutral and aggravating circumstances is observed, while the drink-driving offence, although also showing a linear increase, derives higher mean ratings which only slightly increment across type of circumstance. The drink-driving offence shows lower standard deviations across all types of circumstance compared to the speeding offence suggesting less variation in ratings of seriousness. In general, there is strong agreement amongst participants in their ratings of seriousness given the offence and accompanying circumstances.

A two-way (2 x 3) mixed ANOVA was used to statistically test the hypotheses (data analysis was undertaken using SPSS version 15). As predicted a significant effect was observed for circumstances ($F_{(2,36)} = 91.427$, $p < 0.001$) and for violation ($F_{(1,18)} = 24.121$, $p < 0.001$). A significant interaction between circumstances and

violation was also found (F (2,36) = 39.829, p < 0.001).

For descriptive statistics (see Appendix C3 Experiment 3: Descriptive Statistics) and for the SPSS output (see Appendix D3 Experiment 3: SPSS Output).

4.4 Experiment 4: Rationale

Experiment 3 has shown a significant difference in the ratings of seriousness across three types of circumstance for two motoring transgressions. The following experiment employed an inference task in which participants were presented with the scale of a motoring violation and the circumstances, either aggravating or mitigating, associated with the transgression. The number of mitigating and aggravating circumstances were reduced to eight each with the neutral circumstances excluded. The task for the participant was to judge whether the motorist committing the violation should be fined. The judgement was made following the presentation of all relevant information, that is, *end-of-series* (eos) (cf. Hogarth & Einhorn, 1992). Three predictions were made for the speeding and drink-driving violations: A major offence would lead to greater levels of fine compared to a minor offence and aggravating circumstances would lead to greater rates of fine than mitigating circumstances. Given the observed difference between the two motoring violations in Experiment 3 it was further predicted that drink-driving would be awarded significantly greater fines than the speeding transgression. Formally, the hypotheses can be stated thus:

1. A major offence will receive significantly higher fines than a minor offence for both speeding and drink-driving.

2. Aggravating circumstances will receive significantly higher fines than mitigating circumstances for both speeding and drink-driving violations.

3. An offence of drink-driving will receive significantly higher fines than that of speeding.

4.5 Experiment 4: Method

Participants

Sixty undergraduate students attending the University of Wolverhampton acted as participants. In total 36 females and 21 males participated with three individuals declining to give their gender. The age range was 18-50+ years (median 22 years) with three participants declining to give their age and one participant describing their age as 50+.

Each participant was assigned to one of six experimental conditions with a total of 10 participants per condition. The first condition was a control condition involving a speeding violation and 5 females and 4 males participated. Their ages ranged from 19-33 years (median age 28 years) with one participant declining to give their gender or age. The second condition involved a minor speeding offence and comprised of 5 females and 4 males wit an age range of 18-31 years (median age 21 years). One participant did not provide details of gender or age. Condition three involved a major speeding violation and had 4 females and 6 males as participants, with an age range of 19-39 years (median age 22 years). The fourth condition was a control for a drink-drive violation and had 9 female participants with an age range of 18-42 years (median 19 years). One participant failed to provide details of gender or age. A minor drink-driving offence acted as the fifth condition which included 9 female participants and 1 male participant. Their age range was 18-37 years (median 23.5 years). The final condition consisted of a major drink-driving offence and included 4 female and 6

male participants with an age range of 20-50+ years (median 24.5 years).

Materials

The materials comprised of six separate booklets, one for each of the four

experimental conditions and two control conditions. The booklets were of A4 size and

included an instruction sheet with an example of the task. The example was similar

for all conditions, except the offence related to either speeding or drink-driving. The

example for speeding will suffice to illustrate the given example:

STATEMENTS

**1. If a car driver travels above the legal speed limit and is stopped by the police,
then s/he is liable to a fine.**

2. A car driver travels above the legal speed limit and is stopped by the police.

3. The driver is traveling down a narrow lane.

The following page comprised of a given statement regards either a speeding or drink-

driving offence to which the participant was required to respond. Subsequent pages

comprised of the same motoring transgression but also included an additional

circumstance. Eight mitigating and eight aggravating circumstances were employed

and it was the task of the participant to judge whether the motorist should be fined

given the offence and the circumstance. The judgement was made using a nine point

scale with all 16 circumstances requiring a response.

For the control conditions it was stated that the driver travels above the legal limit

(30mph) or has drunk above the legal limit (2.5 pints) but the specific speed or

amount drunk is not given. However, for the minor and major conditions the actual speed and amount of alcohol consumed was specified. For the minor speeding violation the driver was stated as traveling at 35mph in a 30mph zone. The equivalent minor drink-driving offence specified that the driver had drunk 2.5 pints of alcohol, when the legal limit is 2 pints. The corresponding major offences had the driver traveling at 60mph for the speeding violation and having drunk 5 pints for the drink-driving transgression.

The circumstances used in the experiment are listed below:

Mitigating	**Aggravating**
Late for plane	Late for party
Experienced driver	15 year-old joy-rider
No convictions	Disqualified
2 am	5 pm
Visibility good	Visibility poor
Road empty	Rain & fog
Doctor on-call	Driver in accident
Wife in labour	Teacher

See Appendix A4 (Experiment 4: Materials) for a copy of the booklets used in each of the six conditions.

Design

A mixed experimental design was employed comprising of three factors; type of violation, scale of violation, and circumstances. Type of violation was a between-participants factor with two levels (speeding and drink-driving); scale of violation was also a between-participants factor and had three levels (minor, major, and control); circumstances was a within-participants factor with three levels (mitigating, aggravating and baseline). Three independent variables (I.V.'s) were manipulated, type of violation, scale of violation, and additional circumstances. The dependent

variable (D.V.) was the rating of whether the driver should be fined selected from a 9-point scale.

Procedure

A total of 60 participants were tested in small groups or individually with 10 participants recruited for each of the six conditions. Each participant received one of the six booklets at random and was asked to read the instruction sheet. The opportunity was given for participants to ask questions in order to clarify the nature of the task.

The initial response was for participants to rate whether the offence (speeding or drink-driving) should result in a fine, prior to the introduction of mitigating or aggravating circumstances. This response acted as the baseline condition for the circumstances factor, allowing for a comparison with the mitigating and aggravating circumstances.

Each participant considered only one type of offence (speeding or drink-driving) and only one level of scale of violation (minor, major, or control) but rated the transgression in terms of all 16 additional circumstances.

Participants made their judgements using a 9-point scale. Responses were made by ticking a box beside the response option. More than one response was acceptable, however, participants were requested to indicate their order of preference by numbering their responses (1 = first, 2 = second etc.).

The list of options as presented in the materials is shown below:

(i) the driver should be fined

(ii) the driver ought to be fined

(iii) the driver may be fined

(iv) the driver must be fined

(v) the driver should not be fined

(vi) the driver ought not be fined

(vii) the driver may not be fined

(viii) the driver must not be fined

(ix) none of the above

If option *(ix) none of the above* was chosen, the participant was asked to provide an explanation. No time limit was set to complete the booklet.

The scoring of the responses was as follows, with the participants first choice taken as their score when more than one response was selected:

Response	Score
the driver must be fined	2
the driver should be fined	3
the driver ought to be fined	4
the driver may be fined	5
the driver may not be fined	6
the driver ought not be fined	7
the driver should not be fined	8
the driver must not be fined	9

none of the above   Penalty > fine 1; Penalty < fine 10

The list of responses in the materials differed from the scoring scale to avoid leading

the participant to a particular judgement and to avoid a response bias. For those

responses of *none of the above,* the explanation was considered; a score of 1 was

assigned if the statement suggested the driver should be more than fined (e.g. a

number of participants proposed a custodial sentence) and for those statements

indicating something less than a fine a score of 10 was given (e.g. a police escort was

proposed by more than one participant). If a suggested penalty did not fit the scoring

system a value of zero (0) was assigned, which occurred with 8 responses in the

overall data set.

4.6 Experiment 4: Results

For raw data see Appendix B4 (Experiment 4: Raw Data). All participants completed

the task satisfactorily and all data was included for analysis except for the eight

occasions when the explanation given could not be scored.

Table 4.2 shows the mean and standard deviation scores for the two motoring

violations across, mitigating, aggravating, and baseline circumstances and minor,

major and control conditions. There are several noticeable findings; major offence is

judged more serious than a minor offence across all conditions; the control conditions

in which the scale of violation is not specified tend to fall between the major and

minor conditions, however, this is not consistent; the baseline conditions in which

circumstances are not included consistently score between mitigating and aggravating

circumstances, as shown in Figures 4.1 and 4.2, the only exception being the drink-

driving condition with a major offence.

	BASELINE (no circumstances)		
	Control	Minor Offence	Major Offence
Drink-driving	3.2 (1.69)	2.9 (1.29)	2.8 (2.30)
Speeding	3.63 (1.16)	5.4 (1.78)	3.6 (1.17)
	MITIGATING CICUMSTANCES		
	Control	Minor Offence	Major Offence
Drink-driving	3.8 (1.96)	3.5 (0.90)	2.4 (1.21)
Speeding	5.1 (1.03)	6.0 (1.44)	4.9 (1.55)
	AGGRAVATING CIRCUMSTANCES		
	Control	Minor Offence	Major Offence
Drink-driving	2.5 (1.09)	2.6 (0.77)	1.8 (0.83)
Speeding	3.0 (0.68)	4.1 (0.93)	2.7 (0.65)

Table 4.2: The mean scores and standard deviations (in parentheses) for the two motoring offences, Drink-driving & Speeding, across violation (control, minor offence & major offence) and circumstance (baseline, mitigating & aggravating). The scores show ratings of seriousness; the lower the score the more serious the offence is considered

The drink-driving offence with aggravating circumstances is considered the most serious, as observed by the lowest score. That is, on average participants believe the motorist *must* or *should* be fined. Mitigation with a drink-driving offence is observed to a degree; participants believe the motorist *should* or *ought to* be fined in the control and minor offence conditions. Mitigation does not, however, occur with the major offence. Baseline ratings, where no circumstances are given, tend to fall between aggravating and mitigating circumstances, with participants suggesting the drink-driving motorist *should* be fined and the speeding motorist *ought to* or *may* be fined. The speeding offence is considered more serious when coupled with aggravating

circumstances; participants believing the motorist *should* or *ought to* be fined. A speeding offence with mitigating circumstances is considered the least serious of all presented offences with participants proposing the motorist *may* or *may not* be fined. It is interesting to note that the majority of mean scores are relatively low. There were individual participants who stated that the motorist *should not* or *must not* be fined, however, on average, most offences were considered serious and participants suggested the motorist should receive a fine.
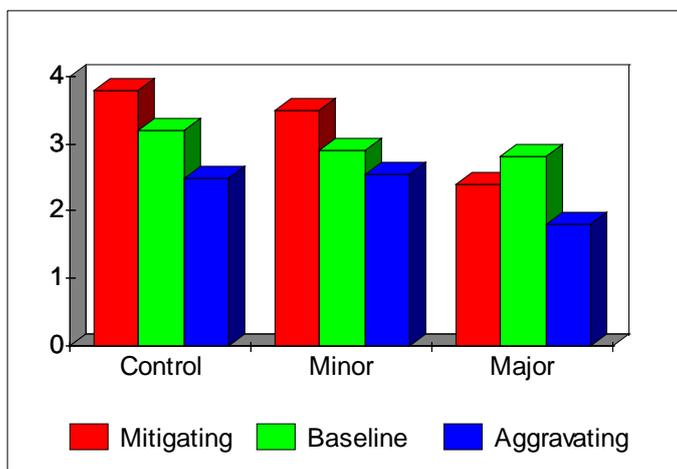
Figure 4.1: Mean scores for Drink-driving across circumstances (baseline, mitigating & aggravating) and violation (control, minor & major)
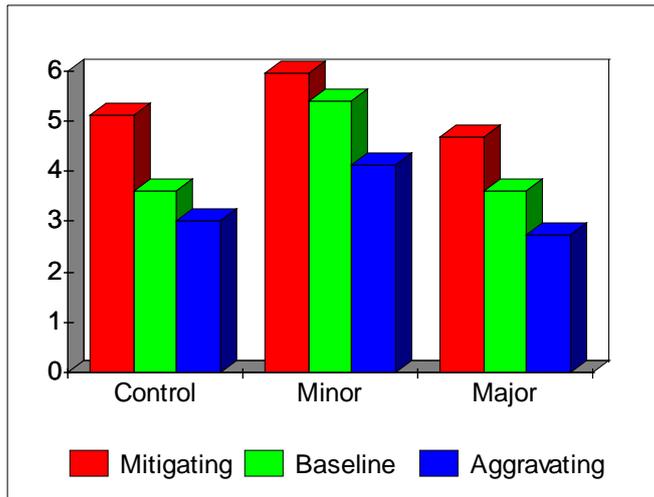
Figure 4.2: Mean scores for Speeding across circumstances (baseline, mitigating & aggravating) and violation (control, minor & major)

Statistical analysis of the data was undertaken using a 3-way (3 x 2 x 3) mixed analysis of variance (ANOVA). Data analysis was undertaken using SPSS version 15. For the within-participants factor of circumstances Mauchly's Test of Sphericity was found to be significant indicating that the principle of homogeneity of variance had been violated; the values from the Greenhouse-Geisser test is, therefore, given. Circumstances showed a significant difference ($F_{(2,87)} = 36.367$, $p < 0.001$); the circumstances x violation interaction was also found to be significant ($F_{(2,87)} = 5.982$, $p = 0.006$). No significant effects were found for circumstances x scale interaction ($F_{(3,87)} = 0.902$, $p = 0.452$) and circumstances x violation x scale ($F_{(3,87)} = 1.2$, $p = 0.315$). For the between-participant factors violation was significant ($F_{(1,52)} = 25.055$, $p < 0.001$) and scale of violation was significant ($F_{(2,52)} = 4.565$), $p = 0.015$) with no interaction effect between violation x scale ($F_{(2,52)} = 1.705$, $p = 0.192$).

4.7 Discussion

The hypotheses for Experiment 3 are supported; significant differences were found across scale of violation (major / minor / control), circumstances (mitigating / aggravating / baseline) and type of motoring offence (speeding / drink-driving).

The experiment shows clear differences in the participants' perception of the two offences. Drink-driving is considered much more serious than speeding, major offences are considered more serious than minor offences and circumstances can either mitigate or aggravate a motoring violation. These results are generally as predicted with the exception being the contrasting finding between the two forms of transgression. As shown in the mean scores presented in Table 4.2 the speeding violation shows a consistent trend of minor, control and major in terms of lowest to highest rating of fine. The control condition is most similar to the major offence, suggesting that when no speed is given (control condition) participants assume a major violation has occurred. The trend is less consistent with the drink-driving offence, in this case the control condition is more similar to the minor condition, across all circumstances, suggesting that without the scale of offence participants assume a minor violation has been committed. Presumably, the participants believe that few drivers would deliberately flout today's strict drink-driving laws but may risk drinking slightly over the legal limit.

The baseline conditions, in which no circumstances are stated, allow a comparison with the conditions incorporating mitigating and aggravating circumstances. The mean scores (see Table 4.2) show a consistent trend in the judgement of a fine, from mitigating, baseline to aggravating. This is as predicted; mitigation reduces the judged

fine in relation to the baseline, while aggravation increases it. Similar observations are made with the drink-driving scenario except for the major offence. Overall, these findings show the modifying effect the inclusion of circumstances has on ones reasoning.

Current theories of deontic reasoning cannot fully accommodate the findings of this experiment. For instance, Pragmatic Reasoning Schema (PRS) theory (Cheng & Holyoak, 1985) suggests all permissions and obligations are the same. This is clearly not the case as it was found here that drink-driving was never condoned, although one could say it was mitigated, even when a minor offence had been committed. Cosmides (1989) may be able to explain the evolutionary advantage of detecting a speeding motorist or drink-driver in terms of taking a prohibited benefit but has greater difficulty explaining how circumstances modify the perception of a *cheater*. Clearly, all *cheaters* are not the same; a drink-driver may be excused if they happen to be a doctor on an emergency call. Mental models theory (Johnson-Laird & Byrne, 1991) and adaptations of it (e.g. Manktelow & Over, 1995) would also require revision to accommodate the findings presented above.

A further difficulty for mental models theory, and possibly other theories as well, is the mean ratings of seriousness for the control groups (as noted above). If individuals are using mental tokens, as mental models theory suggests, then how can participants assign a value to these tokens as these set of results has found? Weightings may be needed to accommodate such findings.

If scale of violation is considered a probabilistic factor then a theory such as

Information Gain (Oaksford & Chater, 1994) or Suppositional Theory (Evans, Handley & Over, 2003) may provide a better explanation. Circumstances may also be explained in terms of Evans & Over's (2004) conditional probability hypothesis; the probability of a motorist receiving a fine is increased when aggravating circumstances are involved and decreased with mitigating circumstances.

There was found to be some variation both between aggravating and mitigating circumstances and within a particular circumstance. Speeding, particularly the minor offence, was readily mitigated. The only circumstances to reliably mitigate in the case of drink-driving were *doctor on call* and *wife in labour*. However, mitigation was not universal, some participants considered a doctor drink-driving as an aggravating circumstance and rated the offence as serious (see Appendix B4 Experiment 4: Raw Data). This can be related to Green, McClelland, Muckli & Simmons (1999) who used a scenario of a mugging in which little sympathy was given to the victim because they had used an alternative route home when the assault occurred.

This experiment supports the findings of Experiment 1 reported in Chapter 3, as scale of violation is again found to be a modifying factor to deontic reasoning. However, this study goes further by demonstrating that not all violations, nor scale of violations, are perceived on equal terms. Circumstances can also interact with an offence altering the perceived seriousness of a violation. These findings once again emphasise that *content* of deontic scenarios must be given as much credence as the *processes* of reasoning.

Mitigating and aggravating circumstances must be added to the growing list of

modifying factors found to influence deontic reasoning. Adding a circumstance to a conditional statement alters the drawn inference in predictable ways, similar to the effects of additional premises (Byrne, 1989; Stevenson & Over, 1995), alternative causes and disabling conditions (Cummins, 1995; Cummins et al., 1991; Fairley et al., 1999), and additional requirements (Fairley et al., 1999). There is also the possibility that motoring offences belong to a superordinate class, such as "transgressions of the law" but that has not been investigated here and must wait for further research. Judgements, in these instances, were made following the presentation of all information, however, in the real world events usually unfold serially and one is required to revise their beliefs as subsequent information comes available. This is to be investigated in the next chapter.

Chapter 5

Judgement Revision with Motoring Violations

5.1 Rationale

The previous experiments reported in this thesis have established that scale of violation, type of transgression and surrounding circumstances all contribute to modifying one's reasoning with deontic conditional statements. They are pragmatic factors as they relate to real world knowledge in familiar social contexts; working rules and motoring transgressions. Experiments 1, 3 and 4 were similar in requiring a judgement to be made after all relevant information had been given. However, in the real world events usually unfold serially allowing for a revision of judgement. The experiment reported in this chapter takes such an approach, thus affording the opportunity to forge an alliance between the fields of reasoning and decision-making.

The proposal that reasoning and decision-making are linked is not new but progress in forging associations between the two fields has been slow. Evans, Over & Manktelow (1993) believe the two fields to be concerned with rationality and suggest that reasoning tasks, particularly deontic reasoning involves the making of decisions. Manktelow & Over (1995) have been particularly forceful in pressing for a decision-theoretic approach to deontic reasoning.

Two fundamental concepts of decision-making are subjective probability and subjective utility and the combination of the two gives subjective expected utility (SEU) theory. Probabilities and utilities are fundamental to reasoning with deontic statements as reported extensively in the literature (see Manktelow, 1999 for a

review). However, subjective expected utility is a normative theory that proposes axioms beyond the cognitive abilities of human reasoning. For this reason, while adopting components of a decision-theoretic approach, Evans et al. (1993) reject the normative account of decision-making.

A simple manipulation to convert a deontic reasoning task into a decision-making task is to present material serially. This method is adopted here and was influenced by Hogarth & Einhorn (1992) who reviewed 60 belief revision studies. Interestingly, in their review only a minority of studies used serial presentation, a response mode they labeled *step-by-step (sbs)*.

Serial (sbs) presentation of material has not been done before using deontic materials, however, there are studies that use a similar procedure. Byrne (1989) and Stevenson & Over (1995) used additional premises, alternative causes and disabling conditions have been manipulated in several studies (Cummins, 1995; Cummins, Lubart, Alksnis, & Rist, 1991; Fairley, Maktelow & Over, 1999) and Fairley et al. (1999) used additional requirements. Any one of these studies could have presented the material step-by-step presumably with the same effects of modifying drawn conclusions.

One advantage of a judgement revision task is that it allows for several independent variables to be manipulated with the same participant. For example, an established finding in the decision-making literature is that costs are weighed more heavily than benefits (Kahneman, Slovic & Tversky, 1982). Similar findings have been obtained with deontic reasoning tasks. For instance, Manktelow & Over (1991) designed a

thematic selection task involving a bingo scenario in which a prize (benefit) did not

produce a facilitation effect. However, when costs are manipulated facilitation occurs

much more readily. Kirby (1994) adapted Griggs & Cox (1982) drink-age rule and

included three manipulations to the standard task. The rule stated, "If a person is

drinking beer, then the person must be over 21 years of age". Participants were cued

to the role of a pub doorman who is given one of the three tasks; don't check as the

management do not want to upset customers and they will be fired if they do so; don't

miss as the management are concerned about illegal drinking and they will be fired if

they miss a guilty person; and check as the management are concerned about illegal

drinking and they will receive a bonus for catching under-aged drinkers. The

conditions involving a cost produced significantly more card selections compared to

the standard task.

The findings with aggravating circumstances and major violations reported thus far in

this thesis may be analogous to costs and mitigating circumstances and minor

offences may have similarities to benefits. Certainly, response patterns show

similarities; aggravating circumstances and major violations produce greater ratings

of seriousness than mitigating circumstances and minor transgressions. A judgement

revision task will allow a more detailed investigation of this proposal.

Presenting information serially also ensures the task is clearly framed. In the context

of reasoning and decision-making, Legrenzi, Girotto & Johnson-Laird (1993) have

called this *focusing*. *Relevance* is of course a related concept (Sperber & Wilson,

1986, 1995) and a judgement revision task would be expected to make the most

relevant information explicit. Inferential reasoning tasks as used throughout this thesis

do not require logical or normative (System 2) thinking; this is a further reason for rejecting the normative SEU theory. Reasoning and judgement revision with tasks involving transgressions of law very much encourage System 1 thinking.

In summary, the experiments reported in Chapters 3 and 4 have established the significance of a number of pragmatic factors, including scale of violation, the nature of an offence and surrounding circumstances. Experiment 4, reported in Chapter 4, demonstrated how pragmatic factors interact to influence an individual's decision. Decisions, in these instances, were made following the presentation of all information i.e. end-of-series (eos); in the real world, however, events usually unfold serially and one is required to revise their beliefs as subsequent information becomes available, that is, step-by-step (sbs) (cf. Hogarth & Einhorn, 1992). The following experiment presents the details of motoring offences in series (sbs) in order to plot the revision of an individual's judgements. The following hypotheses were predicted:

1. Significantly higher fines will be given to a drink-driving violation compared to a speeding offence.

2. A major motoring violation will receive significantly higher fines compared to a minor motoring violation.

3. Motoring violations involving aggravating circumstances will receive significantly higher fines compared to motoring offences with mitigating circumstances.

5.2 Experiment 5: Method

Participants

Forty eight full-time students attending Stourbridge College of Further Education acted as participants. The participants comprised of 37 females and 11 males, with an age range of 16-20 years (median age 17 years).

Materials

The materials comprised of A5 size booklets containing one speeding and one drink-driving scenario. The first page of the booklet provided instructions and an example of the task. The instructions stated:

> *For this task you are required to read the details of two court cases involving motorists. For each motoring offence a series of details will be presented on successive pages. There will be three pages for each case. As the details emerge you will be required to decide which level of fine you would impose on the motorist, based on the information you have been given up to that point: the minimum fine is £0 (no fine) and the maximum fine is £100 (max. fine).*

An example was then given:

> *For example, if you are told that a motorist has been charged with driving with undue care and attention, you may wish to levy a fine of £50, and you would do so by circling the appropriate figure, as shown below:*

£0   £10   £20   £30   £40   £50   £60   £70   £80   £90   £100
(no fine)   *(max. fine)*

The details of each offence were presented over three pages as the type of violation, scale of violation and additional circumstance were presented in-series.

Following the instructions and example, the participant was presented with the type of offence (speeding or drink-driving) and the task of deciding if the motorist should be fined. The rate of fine was determined a scale of £0 (no fine) to £100 (maximum fine) in increments of £10. The subsequent page developed the offence by describing it as a minor or major violation with the participant having the option to revise their initial fine. The next page presented the additional circumstance (either mitigating or aggravating) and the participant was offered the opportunity to revise their rate of fine once again.

An example of the materials is given below:

Rating 0 (R0): A motorist is speeding

Rating 1 (R1): The motorist was traveling at 60 mph in a 30 mph zone

Rating 2 (R2): The motorist is disqualified from driving

Rating 0 (R0) is the initial judgement, rating 1 (R1) refers to the first revised judgement, and Rating 2 (R2) is the second revised judgement.

Having completed one scenario, participants were required to complete a second, contained within the same booklet. The type of motoring offence (speeding / drink-driving) was counter-balanced reducing the possibility of a response bias or learning effect. See Appendix A5 (Experiment 5: Materials) for copies of the materials used in

Experiment 5.

Design

A mixed experimental design was employed with three factors having two levels each; type of violation (speeding and drink-driving), scale of violation (minor and major), and circumstances (mitigating and aggravating). The independent variables (I.V.'s) were type of offence, scale of violation, and circumstances and the dependent variables (D.V.'s) were the three fines (R0, R1, R2).

Procedure

Participants were tested individually or in small groups and assigned to two of the experimental conditions at random. That is, each booklet contained one speeding and one drink-driving scenario, hence, two conditions of the experiment.

The motoring violations were speeding and drink-driving as used in previous experiments. The two levels of scale of violation for speeding were 35mph (minor) and 60mph (major) and for drink-driving the minor offence was 2.5 pints of alcohol and the major violation 5 pints. The number of circumstances were reduced to four; two mitigating - *the motorist's wife is in labour* and *the driver is a doctor on call*, and two aggravating - *the motorist is a 15 year-old joy-rider* and *the motorist is disqualified from driving*. These circumstances were used as they had proven the most powerful in the previous experiment (see Chapter 4).

The task was introduced by having participants read the instructions on the front cover of the booklet and providing the opportunity for any questions. An example of

the task was also provided on the front page. Participants were specifically instructed not to *look-ahead* or turn the pages of the booklet, until they had completed the task on the current page.

The task involved judgement revision with an initial rating and two subsequent revised ratings. The initial rating was made of the type of offence; speeding or drink-driving. The participant was to judge whether the motorist should receive a fine for their transgression using a scale from £0 (no fine) to £100 (maximum fine), incremented in intervals of £10. On completion of this judgement the respondent was requested to turn the page.

The subsequent page presented the violation as either a minor or major offence. The participant was asked to make a second judgement regards fining the motorist given the additional information. Finally, the circumstances, either mitigating or aggravating, associated with the motoring incident were introduced and the respondent made a second revised judgement.

A second motoring scenario was then introduced on the following page of the booklet. The task was as before consisting of an initial judgement, followed by two revised judgements. No time limit was set for the completion of the experiment.

5.3 Experiment 5: Results

All participants completed the task satisfactorily and, therefore, all data was used for analysis. For raw data see Appendix B5 (Experiment 5: Raw Data).

Tables 5.1 and 5.2 show the mean fines and standard deviations levied by participants for each condition of the speeding and drink-driving offence. A number of noticeable differences can be identified; drink-driving receives higher overall fines than speeding; minor offences are given lower fines than major violations for both speeding and drink-driving; and mitigation is shown with both motoring transgressions, although mitigation is condoned to a greater extent with speeding than drink-driving. Scores of standard deviations are high across all judgements, indicating much individual variation between participants.

SPEEDING

	R0	R1	R2
	Speed	35 mph	Aggravating
	42 (12.67)	23 (16.58)	73 (26.05)
	Speed	35 mph	Mitigating
	51 (28.43)	27 (24.62)	18 (17.12)
	Speed	60 mph	Aggravating
	43 (18.64)	57 (21.88)	90 (15.95)
	Speed	60 mph	Mitigating
	47 (22.70)	59 (19.75)	21 (17.82)

Table 5.1: Mean fines and standard deviations (in parentheses) levied against a speeding motorist as the information is presented in series; speeding violation (R0), major or minor offence (R1) and aggravating or mitigating circumstances (R2)

```
-------------------------------------------------------------------------------------------------
DRINK-DRIVING
        R0              R1              R2
        Drink           2.5 pints       Aggravating
        82 (19.46)      59 (25.03)      88 (17.49)

        Drink           2.5 pints       Mitigating
        82 (24.43)      73 (23.09)      43 (42.50)

        Drink           5 pints         Aggravating
        79 (16.21)      83 (9.85)       80 (20.00)

        Drink           5 pints         Mitigating
        83 (15.45)      79 (16.76)      48 (40.86)
-------------------------------------------------------------------------------------------------
```

Table 5.2: Mean fines and standard deviations (in parentheses) levied against a drink-driver as the information is presented in series; drink-driving violation (R0), major or minor offence (R1) and aggravating or mitigating circumstances (R2)

The data was analysed statistically using a three-way (type of violation x scale of violation x circumstance) (2 x 2 x 2) between-participants analysis of variance (ANOVA). Data analysis was undertaken using SPSS version 15.

Three dependent variables were measured corresponding to the initial judgement (R0) and two revised judgements (R1 and R2). We shall consider the findings in terms of each judgement.

For judgement R0, a significant difference was observed between the type of violation, drink-driving and speeding ($F(1,88) = 74.502$, $p < 0.001$).

For judgement R1, scale of violation (major v minor) was significant ($F(1,88) = 34.105$, $p < 0.001$) and scale x violation interaction was also significant ($F(1,88) = 4.907$, $p = 0.029$).

For judgement R2, a significant difference was found between mitigating and

aggravating circumstances ($F_{(1,88)} = 85.235$, $p < 0.001$). A significant interaction was also found between circumstances x violation ($F_{(1,88)} = 4.896$, $p = 0.03$), however, no interaction effect was observed for circumstances x scale ($F_{(1,88)} = 0$, $p = 1.00$). The interaction between violation x scale x circumstances also produced a non-significant result ($F_{(1,88)} = 1.49$, $p = 0.225$).

A series of post hoc t-tests were undertaken to ascertain where judgements were significantly revised. Table 5.3 summarises the results of such an analysis where R0 represents levied fine for offence (speeding or drink-driving), R1 levied fine for scale of violation (major or minor) and R2 levied fine for the circumstance (aggravating or mitigating).

	R0 v R1	R1 v R2
Sp/35/Mit	p<0.01	N.S.
Sp/35/Agg	p<0.01	p<0.001
Sp/60/Mit	N.S.	p<0.01
Sp/60/Agg	p<0.001	p0.01
Dr/2.5/Mit	N.S.	p<0.05
Dr/2.5/Agg	p<0.05	p<0.02
Dr/5/Mit	p<0.05	p<0.05
Dr/5/Agg	p<0.01	N.S.

Table 5.3: Significance findings for post hoc t-tests
R0 v R1 = offence (speeding or drink-driving) versus scale of violation (35/60 mph or 2.5/5 pints)
R1 v R2 = scale of violation versus circumstance (mitigating or aggravating).

The results show that in the majority of cases subsequent information produces a significant revision in a participant's judgement. The major exceptions are R1 v R2 for both Sp/35/Mit and Dr/5/Agg. A possible explanation is that, in the case of the former, R1 (35mph) has a low rate of fine, therefore, a mitigating circumstance (R2)

115

will not significantly alter the rate of fine, that is, there is a floor effect. In the case of the latter, R1 (5 pints) has scored a near maximum rate of fine, and therefore, an aggravating circumstance does not increase the rate of fine, that is, a ceiling effect is observed.

5.4 Discussion

The aim of the experiment reported in this chapter was to investigate judgement revision in serially presented information. The predicted hypotheses have been supported; drink-driving received significantly greater fines than a speeding violation; major offences receive higher fines than minor transgressions; and aggravating circumstances are given higher fines than mitigating circumstances. The post hoc t-tests also produce predictable findings with significant revisions of judgement occurring in most cases. The exceptions may also be accounted for because a speeding offence of 35 mph coupled with mitigating circumstances would be expected to attract lower fines, and conversely a major drink-driving transgression with aggravating circumstances would be expected to attract a high rate of fine. However, there are two inconsistent findings, the revision from a speeding violation to a major (60mph) violation, and drink-driving violation to a minor (2.5 pints) offence. This inconsistency cannot be easily explained and may require further investigation.

Many of the findings are consistent with those of Experiment 4 (see Chapter 4). For example, using the mean scores above it can be observed that the initial judgement of speeding (R0) receives an average fine that falls between the minor and major offences, suggesting that participants are anticipating a serious violation. The major

offence (60 mph) increases the rate of fine but it by no means reaches the maximum level; only when such a violation is coupled with an aggravating circumstance is a near-maximum level of fining observed. With the drink-driving violation the observations are quite different. Much higher rates of fine are observed from the initial judgement of offence; even the minor offence receives relatively high fines and mitigation, although observed, does not occur to the extent observed in the speeding condition. Participants may be able to forgive a driver for speeding while his wife is in labour but find it much more difficult to forgive a driver who has drunk over the legal limit. Similar findings were observed in Experiment 4, although of course the tasks are quite different; one involving a judgement of whether a motorist should be fined and the other requiring a specific fine to be levied.

Caution should be given to some aspects of the experiment and to some findings. The rating scale used a range of fines from £0 - £100. This scale may not be appropriate, particularly when one considers that under current law drink-driving carries an automatic driving ban and may result in a custodial sentence. An alternative rating scale, more consistent with current sentencing, may have been more appropriate. Several significant results reported in Table 5.3 achieve a probability value of <0.05 which may be indicative of a type 1 error. However, these results occur with the drink-driving violation only and in cases were revised judgements may not be expected, for example, minor violation (2.5 pints) with mitigating circumstances. It is likely that in a replicated study the same results would not be observed, however, the finding is interesting because it supports the view that not all violations are perceived the same.

This experiment has also shown how serial presentation of material can successfully transform a deontic reasoning task into a judgement revision task suggesting a similar method could be applied to research with additional premises (e.g. Byrne, 1989) and disabling conditions (e.g. Cummins, 1995, Cummins et al., 1991). It can also be argued that serial presentation ensures the most important information is attended to, that is, *relevance* holds (Sperber & Wilson, 1986, 1995).

Major transgressions and aggravating circumstances may be likened to the utility of a cost as they increase a response (in this case a fine), while minor transgressions and mitigating circumstances decrease a response. Of course transgressions and circumstances are not costs and benefits, in terms of subjective utilities, but they may evoke similar cognitive processes.

The findings of this experiment also support Evans et al.'s (1993) and Oaksford & Chater's (1994) claims for a decision-theoretic approach to reasoning. Scale of violation and associated circumstances may be suitable candidates for probabilistic variables within such an approach, and the results do support this notion. Revised judgements can be considered as revised probabilities.

In terms of a theoretical explanation of the findings a number of theories can be considered. Schema theories, such as Pragmatic Reasoning Schema (PRS) theory (Cheng & Holyoak, 1985) and Social Contract Theory (SCT) (Cosmides, 1989) have not proposed how one's schema may be revised as subsequent information becomes available. Indeed, PRS theory is unlikely to be able to explain judgement revision, other than to propose the evocation of an alternative permission or obligation schema.

118

Cosmides' SCT theory suggests a cheater detection algorithm is evoked when a social contract has been violated, however, at what point does the Darwinian algorithm become operational in a judgement revision task involving serially presented material. It is possible that major and minor violations and aggravating and mitigating circumstances could be conceived of as costs and benefits (as discussed above), however, the theory would require substantial amendment to accommodate a range of pragmatic factors.

Mental models theory (MMT) (Johnson-Laird & Byrne, 1991) would also require modification to accommodate judgement revision and the pragmatics of violations. Indeed Johnson-Laird & Byrne (2002) have included a component to deal with context, which they call *pragmatic modulation*. This allows potentially inconsistent or fallacious models to be rejected. However, this adapted version of MMT has its critics for lacking psychological evidence (Evans, Over, & Handley, 2005). It is unlikely that the revised mental models theory would be able to explain the findings reported in this thesis as the research of Johnson-Laird and colleagues rely invariably on abstract and arbitrary content.

A more appropriate explanation of the findings reported in Experiment 5 is to assume that judgement revision in these familiar contexts involves System 1 thinking and the revision of one's decision involves revised probabilities, as suggested in the suppositional theory of Evans and colleagues (Evans, Handley & Over, 2003; Evans Over & Handley, 2003). With revised judgements, the P(q|p) would change as subsequent information is made available. For instance, knowing a motorist drives above the legal speed limit (p) and is liable to a fine (q) indicates the probability of

the motorist being fined, given they travel above the legal limit. Given further information regards actual speed and surrounding circumstances, judgements are revised on the basis that the probabilities have changed; major offence and aggravating circumstances will increase the probability of a fine and a minor offence and mitigating circumstances decrease the probability. This theory offers a more parsimonious explanation and can account for many of the findings reported in this and previous chapters. However, the precise mechanism for integrating pragmatic context and framing the relevant content is not specified in the theory.

Chapter 6

Power of Source as a Pragmatic Factor of Deontic Reasoning

6.1 Rationale

Thus far, scale of violation and mitigating and aggravating circumstances have been

shown to be influential pragmatic factors within the deontic contexts of a working

rule and motoring violations. These effects have been demonstrated with inferential

reasoning tasks and a task requiring judgement revision. A further pragmatic factor is

to be explored in this chapter, that of power of source. Mitigation and aggravation

may be universal concepts, applying to everyone. However, power or authority is

associated with ones role or rank and is a hierarchical concept. Only if you have a

particular role or position with responsibility can you have power. Power is therefore

not universal but is held in degrees. In short, power of source may be an additional

pragmatic factor but one quite distinct from those covered thus far in this thesis.

It seems reasonable to expect that the power or authority of an individual making a

conditional statement or conveying information will determine the perceived felicity

of that statement and whether a given action will follow. Indeed, there is evidence

from previous research for the effectiveness of power as a pragmatic factor (Ray,

Reynolds & Carranza, 1989). However, the emphasis of Ray et al.'s research was on

the reasoning of disjunctives with the aim of establishing whether the preferred

reading was inclusive or exclusive. Thus, the influence of power was secondary and

the results were inconclusive, which may have been due to the authority figures used

in the scenarios. For example, one source of power was a shop assistant which may be

a case where the felicity is brought into question. A shop assistant may not be

perceived as holding sufficient authority to assert a felicitous inducement and make their statement true.

Power is not a unitary concept as Raven (French & Raven, 1959, Raven, 1965) realised by suggesting six types of power – reward, coercive, informational, expert, legitimate, and referent. A number of these power types are incorporated in the scenarios used in the experiments reported in this chapter. For instance, a medical consultant would have *expert* power, parents use *reward* as a form of power over their children, and even those of lower status may exert *coercive* or *informational* power, such as the receptionist who prevents you seeing the bank manager.

Power has been explored extensively by researchers in the field of social cognition. For example, Bochner & Insko (1966) found a message from a Nobel Prize winner was more influential than the same message from a less prestigious source. Social psychological studies have shown the detrimental effects of low power status. Groups, such as children (De Paulo & Coleman, 1986, 1987), low-ranking individuals and historically women (Hacker, 1951) have served low ranking positions with little control over the outcomes of others. Keltner and colleagues have considered how one's relative power influences exhibited behaviour with those of higher status having increased rewards and freedom, resulting in approach-related tendencies while those with reduced power show inhibited behavioural tendencies (Keltner, Gruenfeld & Anderson, 2003).

Fewer studies have been reported in the thinking and reasoning literature, however, similar results have been reported. For instance, Stevenson & Over (2001) explored

conditional reasoning and found experts were more likely to be believed and to introduce uncertainty into an argument than novices. The first experiment reported in this chapter uses conditional statements within deontic contexts.

Power has been the central theme of at least one theory of deontic reasoning. Cummins (1999, 2000) considers power and dominance relations as a driving force of human evolution which is observed in social hierarchies. Cheater detection, as proposed by social exchange theory (Cosmides, 1989), is more likely to be observed in individuals of low status.

However, the theory of dominance hierarchies has been questioned (Chater & Oaksford, 1996) and a probabilistic account of power may be more appropriate. Evans, Handley & Over (2003) reviewed probabilistic accounts of conditional reasoning and suggested conditional statements of the form, if p then q, can be considered in terms of degrees of belief and represented as the probability of the consequent given the antecedent ($P(q \mid p)$). Historically, probabilistic reasoning has been investigated using abstract materials (e.g. Tversky & Kahneman, 1973) and this trend continues to the present day (see for example, Evans et al., 2003; Oaksford, Chater & Larkin, 2000; and Oberauer & Wilhelm, 2003). However, a number of studies have investigated the pragmatic richness of everyday reasoning using thematic materials (Oaksford et al., 2000; Stevenson & Over, 1995), realistic inducements (Evans & Twyman-Musgrove, 1998; Newstead, Griggs & Chrostowski, 1984) and deontic conditionals (Over, Manktelow & Hadjichristidis, 2004).

The power or authority of the speaker (agent) making a conditional deontic statement has not been manipulated before although related studies do exist (e.g. Stevenson & Over, 2001; Ray et al., 1989). The first experiment to be reported in this chapter investigated the power of a message source with conditional statements across two forms of scenario – informal (family and medical) and formal (armed forces & employment). It is predicted that a significant trend will be observed in relation to power and the likelihood of action. In all cases an individual of higher social rank is predicted to be rated as more likely to make a given conditional statement true than an individual of lower social rank. A further prediction can be made regards the form of scenario; a formal scenario in which rank, authority and power of its members is institutionalised may show greater effects of power than informal scenarios where institutionalised authority is absent.

The hypotheses can be stated formally as follows:

1. A significant effect of power will be observed across individuals of high, medium and low power. Further, the predicted difference will demonstrate a significant trend.

2. A significant difference will be observed in likelihood ratings of action between formal and informal power scenarios.

6.2 Pilot Study: Method

Prior to experimentation a pilot study was undertaken to obtain scaled power relations within a number of social contexts.

Participants

Twenty undergraduate students attending the University of Wolverhampton acted as independent judges rating the power or authority of 21 individuals within seven social contexts. No details regards gender or age of judges was recorded.

Materials

A three page booklet containing instructions with an example followed by 21 individual roles were used as the materials. The instructions introduced the task which was to rate individual roles within seven social contexts for power or authority on a 5-point scale. The example used a Grandmother and showed a rating of *medium* power with the statement,

*If you were asked to rate the **Power** (**influence** or **authority**) of a Grandmother and you believed she had a medium degree of power, then you would tick the **medium** box, as shown below:*

Please rate the POWER of a <u>Grandmother</u>:

High		*Medium*		*Low*
[ ]	*[ ]*	*[✔]*	*[ ]*	*[ ]*

The seven social scenarios were: Family, Education, Armed Forces, Medical, Employment, Football Team, and Legal. Three individual roles associated with each

scenario were chosen with some consideration given to the potential hierarchical power. For example, the Family scenario had the roles of Father, Uncle and Brother. See Appendix A6a (Pilot Study: Materials) for a copy of the booklet used in this pilot study.

Design

A within-participants design was used with all participants completing all ratings of individual authority within seven social contexts.

Procedure

Participants were tested individually or in small groups. After consenting to participate in the ratings study participants were asked to read the instructions and example. Time was given for any questions to clarify the nature of the task. The task involved rating individual roles for power or authority on a 5-point scale, where 5 represented high power/authority and 1 low power/authority. The 21 individual roles were presented randomly within the booklet and no mention of the social categories was made to participants. All 21 roles were judged using the same scale and no time limit was set for the completion of the task.

The ratings of each role by the 20 participants was totalled giving a minimum possible score of 20 and a maximum possible rating of 100.

Results

Table 6.1 shows the total ratings of power for each individual within each of the seven social contexts.

SCENARIO	ROLE 1	ROLE 2	ROLE 3
FAMILY	FATHER 76	UNCLE* 48	BROTHER 56
EDUCATION	HEAD TEACHER 88	PREFECT 51	FELLOW PUPIL 49
ARMED FORCES	SERGEANT 82	CORPORAL 75	CADET 36
MEDICAL	CONSULTANT 87	G.P. 81	STRANGER** 41
EMPLOYMENT	MANAGING DIRECTOR 87	SUPERVISOR 77	COLLEAGUE 53
FOOTBALL TEAM	MANAGER 81	CAPTAIN 70	TEAM PLAYER 51
LEGAL	SOLICITOR 82	LEGAL SECRETARY 57	FRIEND 71

Table 6.1: The total ratings of power for individuals for seven scenarios. Twenty participants (N=20) were asked to judge all 21 individuals for power on a 5-point scale (ranging from High – Medium – Low). The minimum possible score was 20 and the maximum possible score was 100.

NOTE: * The Uncle was designated as the individual having Medium power.
** Stranger in the medical scenario was replaced with Receptionist as it was considered a more appropriate role for a medical scenario.

From the ratings four scenarios were chosen as the materials for experimentation.

Three individuals were identified as representing high, medium and low power within

informal (family & medical) and formal (armed forces & employment) social

contexts. Formal social contexts differ from informal in respect of institutional power

invested in higher ranking roles; the power associated with informal scenarios is not

institutional. Table 6.2 shows the roles assigned to high, medium and low power to

the scenarios of family & medical (informal) and armed forces & employment

(formal).

Scenario	Power of Source		
	High	**Medium**	**Low**
Family (informal)	Father	Uncle	Brother
Medical (informal)	Consultant	G.P.	Receptionist
Armed Forces (formal)	Sergeant	Corporal	Cadet
Employment (formal)	Managing Director	Supervisor	Colleague

Table 6.2: Assigned power of source, from rating of judges, for high, medium and low power within informal (family & medical) and formal (armed forces & employment) scenarios

6.3 Experiment 6: Method

Participants

Participants comprised of 10 undergraduate students attending the University of

Wolverhampton. Six participants were female and four male with an age range of 19-

35 years (median age 19.5 years). Participants completed all conditions of the

experiment.

Materials

The materials comprised of a single A4 size booklet containing four social scenarios. Two scenarios were classified as formal (Armed Forces and an Employment hierarchy) and a further two as informal (Family and a Medical hierarchy). Power of source was embedded within each scenario using individuals of differing status.

The front page of the booklet provided instructions and an example of the task. The instructions stated,

> *Inside this booklet you will find a series of statements. Each statement involves one person asking another to undertake some action. Given the statement and the person who makes it, you must judge the likelihood that the individual will act.*

The example involved a permission spoken to a young girl,

> *"If you run an errand for me, then you can watch T.V."*

The participant was then asked to consider the likelihood that the girl would run the errand given that the statement had been made by the young girl's mother. The judgement involved ticking one of five options presented in a Likert scale, from Very likely to Very unlikely. For the example, the Likely option was ticked with the statement:

*If you believe it is likely that the young girl will run the errand, you would tick the **likely** box as shown above.*

All statements were to be completed with a single judgement per statement.

On the following four pages, conditional statements were presented for each of the scenario types with participants judging the likelihood of an individual acting, given that a conditional statement was made by individuals of differing status. Each scenario appeared on a separate page with the three levels of power randomized. For example, with the informal family scenario participants were presented with the following statement:

Spoken to a boy: "If it is the weekend, then you may go to the cinema"

Participants had to judge the likelihood of the boy going to the cinema, given that the statement had been made by the boy's father (high power), uncle (medium power) or brother (low power). Judgements were indicated on the Likert scale shown below:

Very likely   [ ]
Likely   [ ]
Unsure   [ ]
Unlikely   [ ]
Very unlikely   [ ]

See Appendix A6b (Experiment 6: Materials) for a copy of the booklet used in this experiment.

Design

A within-participants design was used with two factors; power of source and formality of scenario. Power of source had three levels, high, medium and low while scenario had two levels, formal and informal. Power of source and formality of scenario acted as independent variables (I.V.'s) and ratings of likelihood was the dependent variable (D.V.).

Procedure

The 10 participants completed all four conditions of the experiment and were tested individually. After consenting to taking part in the experiment the participants were asked to read the instructions on the front page of the booklet. The opportunity was given for questions to clarify the nature of the task

The task for the participant was to judge the likelihood of action by one individual given a statement made by another individual within a particular social context. The target statement for the Family scenario involved a conditional permission spoken to a boy:

If it is the weekend, then you may go to the cinema

The participant was informed that it was the weekend and were asked to judge the likelihood of the boy going to the cinema, given that the statement was made by three individuals of varying power; brother, uncle, and father classified as having low, medium and high power status, respectively. Three individual judgements were made on the 5-point Likert scale, one for each level of authority.

The task was repeated for the three additional scenarios. The Medical scenario comprised of the statement:

With your illness, you may only drink alcohol in moderation

The participant was asked to judge the likelihood that the patient would drink alcohol in moderation given that the statement was made by a Receptionist, General Practitioner (GP) or Consultant; these three individuals representing low, medium and high power, respectively.

The Armed Forces scenario used a statement regards a potential recruit:

If you are thinking of joining the army, then you may visit the barracks

The conditional statement was said to be made by either a Cadet, Corporal or Sergeant (representing low-high power of source). The likelihood judgement was made on whether the potential recruit would visit the barracks using the Likert scale.

The final scenario involved an Employment context which used the conditional statement:

If you want to improve your skills, then you may attend a training course

The three levels of authority were Colleague, Supervisor and Managing Director (MD) representing low, medium and high power, respectively. A judgement was

made on whether the employee would attend a training course given the authority of the individual making the statement.

The order in which the scenarios were presented in the booklet was randomised as were the ordering of the power instantiations. The Family and Medical scenarios were classified as *informal* and the Armed Forces and Employment scenarios as *formal*. The latter contexts contain a strict hierarchical structure of power supported by organizational rules or rules of law; such a hierarchy does not exist in more informal contexts. Each scenario included a conditional statement, except for the Medical context. A conditional statement could have been derived, such as:

If you have your illness, then you may only drink alcohol in moderation

However, this statement seemed clumsy and, therefore, the wording was changed to fit a more natural style of spoken prose.

Responses were scored on a 5-point scale; Very likely scored 5, Very unlikely scored 1. No time limit was set for the participant to complete the experiment.

6.4 Results

For raw data see Appendix B6b (Experiment 6: Raw Data). The experiment was completed satisfactorily by all participants and all data was retained for analysis.

The mean likelihood ratings of action and corresponding standard deviations are given in Table 6.3. A consistent trend is observed with the likelihood of action

increasing as a function of power, except for the Armed Forces scenario where the

trend is slightly reversed for the medium and high power sources (Corporal and

Sergeant, respectively).

Scenario	Mean Ratings and standard deviations ( )		
	Low	Med	High
Family	3.2 (0.99)	3.6 (0.95)	4.3 (0.41)
Medical	2.4 (1.17)	4.2 (0.57)	4.8 ().32)
Armed Forces	2.6 (0.84)	4.3 (1.06)	4.2 (0.63)
Employment	3.2 (0.63)	4.2 (0.42)	5.0 (0.00)

Table 6.3: Mean judgement ratings of action and standard deviations (in parentheses) following a conditional statement from a source of low, medium & high power within informal (Family & Medical) and formal (Armed Forces & Employment) scenarios.

The initial inferential statistics conducted on the data comprised of a two-way (2 x 3)

within-participants analysis of variance (ANOVA). The analysis was undertaken

using SPSS (version 15). Mauchly's Test revealed that sphericity had been violated

for the power factor and, therefore, reported values use the Greenhouse-Geisser test.

A significant effect was found for power ($F_{(2,38)} = 51.458$, $p < 0.001$) but not for

scenario ($F_{(1,19)} = 1.717$, $p = 0.206$). A non-significant effect was also found for the

power x scenario interaction ($F_{(2,38)} = 0.496$, $p = 0.613$).

To test the prediction of a trend from low to high power in the likelihood of action a

series of Page's Trend tests were calculated by hand. Significant trends were obtained

for all scenarios: Family, L = 131, $p < 0.01$; Medical, L = 138, $p < 0.01$; Armed

Forces, L = 132.5, $p < 0.01$; Employment, L = 138 (k = 3, n = 10 in all cases).

However, there is a deviation in the trend for Armed Forces as medium and high

power were reversed.

6.5 Experiment 7: Rationale

Experiment 6 established power of source as a pragmatic factor in deontic reasoning contexts, however, the formality of the given context did not significantly differ between formal and informal scenarios. The aim of Experiment 7 was to investigate the generality of power of source as a pragmatic factor of deontic reasoning. To this end, the experiment was broadened in two respects: firstly, the semantic domain was extended from permissions to inducements, specifically promises, threats and warnings and secondly, the syntax of the given statements was generalized from conditional to conjunctive and disjunctive statements.

Much of the research investigating inducements has focused on interpretation. There has been extensive debate, without agreement, as to whether disjunctive statements are preferentially interpreted as inclusive or exclusive (Braine & Rumain, 1981; Hurford, 1974, Lakoff, 1971; Newstead, Griggs & Chrostowski, 1984; Pelletier, 1977). Inducements are used in a radically different way in the current experiment by investigating whether the authority of a given statement has the same influence when expressed as a promise, threat or warning as they appeared to have been when expressed as a conditional in the previous experiment. The extension to disjunctive and conjunctive statements is novel as previous studies of deontic reasoning have almost exclusively concentrated on conditional, if…then, rules.

There is some evidence to support the proposal that power of source can be applied to inducements. In a comprehensive series of experiments Ohm & Thompson (2004) investigated conditional reasoning using realistic materials embedded in inducements. A speaker who was perceived to have control over consequences or outcome was

considered more likely to change the addressee's behaviour. In the scenarios used by Ohm & Thompson many include powerful speakers, including a professor, policeman, judge and parent and also individuals with less status, for instance, a brother and work colleague.

Newstead, Ellis, Evans & Dennis (1997) also considered the control a speaker has over outcomes. They used conditional statements, including causal (e.g. if the lorry is heavier than the legal limit then the alarm bell will ring) and deontic (e.g. if you wash the car then I'll let you borrow it tonight). In a series of experiments judgements of truth and falsity and judgements involving the four conditional inferences (Modus Ponens, Modus Tollens, Denial of the Antecedent and Affirming the Consequent) were taken. The findings demonstrated that deontic conditionals produced higher rates across all inferences compared to non-deontic conditionals. This would suggest that deontic reasoning tasks provide ideal contexts in which power relations will be observed.

Researchers have defined inducements in various ways. Newstead, Ellis, Evans & Dennis (1997) considered promises and threats as both kinds of inducements but considered warnings as a form of advice. For this reason, the inducements used in this experiment are carefully defined. Threats can be represented as a conditional, conjunctive or disjunctive statement; however, disjunctives involve a negative antecedent. Promises can be represented as a conditional and conjunctive statement but not as a disjunctive. Newstead et al. (1984) did present promises as disjunctive statements but they were options between two choices rather than a requirement for action before the promise is fulfilled. A warning may be represented as a

precautionary conditional statement of future negative consequences for the addressee (Fillenbaum, 1976) or alternatively a conditional warning may be spoken from one individual to another stating that, if you p, then negative q will result from an external agent. Warnings are usually hypothetical (Searle, 1969) emphasising precaution on the side of the addressee to a potential negative consequence with the consequence likely to come from an external agent.

Searle (1969) distinguished between sincere and insincere promises. This relates to the felicity of the statement. An insincere promise is one which the speaker does not (or cannot) fulfill. The essential feature of a promise is that it is an obligation to perform a particular act, arguably more so than threats and warnings. In addition, a promise must be non-defective, the thing promised must be something that the promisee wants done, and the promiser must be able to fulfill the offer. The inducement must have a point. There is no sense in offering something that was to be done anyway (this may apply to parents and politicians). The parent who promises their child, "If you are well behaved this week, I will take you on holiday" is making a defective promise if the holiday had been booked months previously. The sincerity and defectiveness of an inducement is particularly relevant to the authority of the individual making that inducement.

Promises, threats and two types of warning were presented within the formal and informal scenarios introduced in Experiment 6. It was hypothesized that ratings of action would follow a trend from high to low power and this prediction was expected to hold across all types of inducement and for each syntactical form. In formal terms, the hypothesis predicted:

1. A significant trend will be observed from low to high power across formal and informal scenarios.

6.6 Experiment 7: Method

Participants

Forty eight undergraduate students attending the University of Wolverhampton acted as participants. Thirty three participants were female and 15 male with an age range of 18-35 years (median age 20 years). Twelve participants were assigned to each of four experimental conditions; Threat, Warning, Promise (agential) and Promise (external).

Materials

The materials comprised of four separate A4 size booklets, relating to four types of inducement; threat, warning, promise (agential) and promise (external). Participants completed one booklet only.

The first page of each booklet provided instructions and an example. The instruction read,

> *Inside this booklet you will find a series of statements. Each statement involves one person warning / threatening / promising another* (the appropriate inducement was inserted for each condition).

The example scenario used in all conditions involved a young girl and homework but the wording differed to match the context of the inducement. The statement for each of the inducement types used in the task examples is given below:

Threat: Finish your homework, or I will punish you

Warning: Finish your homework, or your teacher may punish you

Promise (agential): Finish your homework and I will give you a treat

Promise (external): Finish your homework and your teacher will give you a treat

The instructions then stated:

Given that the [inducement] is made by the young girl's <u>Mother</u>, how likely is it that the girl will finish the homework?

A 5-point Likert scale was shown with the Likely selection checked (see below).

Very likely   [ ]
Likely   [✓]
Unsure   [ ]
Unlikely   [ ]
Very unlikely   [ ]

Hence, the task for the participant was to judge the likelihood of action, given the source of the inducement.

The following pages introduced one of the four inducements, presented as a statement spoken by one individual to another. Threats and warnings were stated as a

disjunctive ('or' statements) while the two forms of promise were stated as conjunctives ('and' statements).

The four types of inducement were Threat, Warning, Promise (agential) and Promise (external). Each inducement was set in a series of social contexts containing three levels of power – high, medium and low. The social scenarios and levels of power included: Family (Father, Uncle, Brother), Medical (Consultant, GP, Nurse), Armed Forces (Sergeant, Captain, Cadet), Employment (Managing Director, Supervisor, Colleague).

Using the family scenario as an example each of the four conditions can be illustrated. For the conjunctive promise (external source) condition the participant was given the following statement:

> **A boy is promised: "Mow the lawn and your mother will give you your pocket money"**

Three measures were then taken of the likelihood of the boy mowing the lawn given that the statement was made by the Father, Uncle and Brother of the boy.
The statements for the further three conditions were as follows:

Conjunctive promise (agential source) condition:

> **A boy is promised: "Mow the lawn and I will give you your pocket money"**

Disjunctive warning condition:

> **A boy is warned: "Mow the lawn or your mother may not give you your pocket money"**

Disjunctive threat condition:

A boy is threatened: "Mow the lawn or I will not give you your pocket money"

See Appendix A7 (Experiment 7: Materials) for the full set of materials used in this experiment.

Design

A mixed experimental design was employed with two factors; type of inducement and power of source. The inducement type comprised four levels – threat, warning, promise (agential) and promise (external) and was a between-participants factor. Power of source had three levels – low, medium, and high and was a within-participants factor.

Procedure

Each of the four inducement types was operationally defined as follows:

Warning: a statement uttered by an agent to a receiver in which a negative consequence is asserted should a specific action be followed; the consequence being administered by an external agent.

Threat: a statement uttered by an agent to a receiver in which a negative consequence is asserted should a specific action be followed; the consequence being administered by the agent.

Promise (agential): a statement uttered by an agent to a receiver in which the consequence is positive should a specific action be followed; the consequence being delivered by the agent.

Promise (external): a statement uttered by an agent to a receiver in which the consequence is positive should a specific action be followed; the consequence being delivered by a third party.

Using the family scenario as an example each of the four conditions can be illustrated. For the conjunctive promise (external source) condition the participant was given the following statement:

A boy is promised: "Mow the lawn and your mother will give you your pocket money"

Three measures were then taken of the likelihood of the boy mowing the lawn given that the statement was made by the Father, Uncle and Brother of the boy.

The statements for the further three conditions were as follows:

Conjunctive promise (agential source) condition:

A boy is promised: "Mow the lawn and I will give you your pocket money"

Disjunctive warning condition:

A boy is warned: "Mow the lawn or your mother may not give you your pocket money"

Disjunctive threat condition:

A boy is threatened: "Mow the lawn or I will not give you your pocket money"

Inducements may be stated using alternative conjunctive and disjunctive statements. For instance, a warning or threat may be stated as a conjunctive by changing the syntax of the statement. A boy may be warned, "Don't mow the lawn **and** your mother may not give you your pocket money". However, the statement seems rather clumsy compared to the disjunctive form. Similarly with threats, such as, "Don't mow the lawn and I will not give you your pocket money". The double negatives are most likely to result in the statement being transformed to an active conditional clause ("Mow the lawn and I will give you your pocket money"). In both cases, the disjunctive appears the more natural form of expression and was, therefore adopted for this experiment. Promises are not naturally stated as disjunctives. As mentioned above Newstead et al. (1984) devised disjunctive promises such as, "We will either go shopping on Friday or have a meal together on Saturday". This statement was couched in a scenario involving a promise but actually involves a choice between two positive actions. This kind of disjunctive choice was not adopted in the current experiment and, therefore, promises were confined to conjunctive statements only.

Participants judged likelihood of action on the same Likert scale as employed in the previous experiment (see Section 6.2), with very unlikely scoring 1 and very likely scoring 5. Thus a high score indicated expected action given the inducement of warning, threat or promise.

6.7 Results

For raw data see Appendix B7 (Experiment 7: Raw Data). All participants completed

the experiment satisfactorily and all data was used for analysis.

Mean ratings for likelihood of action for each type of inducement and level of power

are given in Table 6.4 below. The mean likelihood ratings follow the predicted

directions for all cases except the armed forces scenario within the disjunctive

warning condition. For all cases there is a substantial difference in the likelihood

ratings between the low and medium power of source, however, when comparing the

medium and high power sources the difference is less pronounced, particularly for the

armed forces scenario. This may reflect the respondent's uncertainty regards the

commands of hierarchy within the armed forces, as reported in the previous

experiment. Overall, however, the ratings follow those found in Experiment 6 and

suggest that the influence of power with conditional statements can be applied to

inducements expressed as conjunctive or disjunctive statements.

Disjunctive Threats

Power of Source

	Low	Med	High
Family	2.17 (0.63)	3.75 (0.45)	4.75 (0.45)
Medical	2.75 (1.06)	3.75 (0.97)	4.25 (0.62)
Armed forces	2.42 (1.24)	4.42 (0.65)	4.67 (0.49)
Employment	2.00 (0.60)	4.00 (0.85)	4.83 (0.39)

Disjunctive Warnings

Power of Source

	Low	Med	High
Family	2.25 (0.97)	3.08 (1.31)	4.58 (0.67)
Medical	2.50 (1.17)	3.50 (1.17)	4.17 (0.72)
Armed forces	3.33 (0.78)	4.67 (0.49)	4.50 (0.52)
Employment	1.92 (0.54)	3.92 (0.79)	5.00 (0.00)

Conjunctive Promises (agential)

Power of Source

	Low	Med	High
Family	1.92 (0.90)	4.00 (0.95)	4.42 (0.51)
Medical	2.67 (0.89)	3.08 (1.00)	3.42 (1.08)
Armed forces	2.58 (1.00)	4.33 (0.65)	4.42 (0.67)
Employment	1.42 (0.67)	4.08 (0.51)	4.75 (0.45)

Conjunctive Promises (external)

Power of Source

	Low	Med	High
Family	1.75 (0.87)	2.75 (1.22)	4.33 (0.49)
Medical	2.67 (1.23)	3.17 (1.53)	3.67 (1.23)
Armed forces	2.75 (0.97)	4.00 (0.74)	4.08 (1.16)
Employment	2.75 (0.87)	3.92 (1.08)	4.83 (0.39)

Table 6.4: Mean likelihood ratings and standard deviations (in parentheses) for different levels of power of source (low, medium, high) for the four different statement types (threat, warning, promise agential & promise external) and four scenarios (family, medical, armed forces & employment)

The statistical analysis of the data comprised of a series of Page's trend tests, conducted by hand, across scenarios for each inducement type. In all cases significant findings were obtained; Family Threat, $L = 166$, $p < 0.001$; Medical Threat, $L = 158.5$, $p < 0.001$, Armed Forces Threat, $L = 160.5$, $p < 0.001$; Employment Threat, $L = 166$, $p < 0.001$; Family Warning, $L = 163.5$, $p < 0.001$; Medical Warning, $L = 160.5$, $p < 0.001$, Armed Forces Threat, $L = 158.5$, $p < 0.01$; Employment Threat, $L = 166.5$, $p < 0.001$; Family Promise (agential), $L = 164$, $p < 0.001$; Medical Promise (agential), $L = 153.5$, $p < 0.05$, Armed Forces Promise (agential), $L = 161$, $p < 0.001$; Employment Promise (agential), $L = 166$, $p < 0.001$; Family Promise (external), $L = 166.5$, $p < 0.001$; Medical Promise (agential), $L = 154.5$, $p < 0.05$, Armed Forces Promise (agential), $L = 156.5$, $p < 0.01$; Employment Promise (agential), $L = 158.5$, $p < 0.01$; $k = 3$, $n = 12$ in all cases.

In only the armed forces scenario is the predicted trend violated, occurring with the inducements of both warning and threat. These findings demonstrate consistent significant trends in likelihood ratings across the levels of power and further establishes the authority of a message source as a prominent pragmatic factor of deontic reasoning.

6.8 Experiment 8: Rationale

Experiment 8 develops the previous two experiments by exploring the hypotheses that a statement from a high power source is both more credible and has a higher probability of being made true than a statement from a low power source. A high power source is more likely than a low power source to be able to grant a benefit for compliance or levy a sanction in the case of transgression. Further, an individual of higher ranking may also be perceived as offering a more felicitous message, which may apply particularly to warnings and external promises where rewards and sanctions are outside of their individual control but where their perceived knowledge and wisdom allow them to more accurately predict the actions of others. For these reasons, direct control of rewards and sanctions (in the case of threats and agential promises) and perceived wisdom in predicting the actions of others (in the case of warnings and external promises), one can hypothesise that a message from a high power source will be rated as more credible and be more likely to make the consequent of their statements true compared to a source of lower rank.

Thus far it has been speculated that the findings with pragmatic factors of deontic reasoning may be explained by the probabilistic theory proposed by Evans and colleagues (Evans, Handley & Over, 2003; Evans, Over & Handley, 2003; Evans & Over, 2004). The next experiment provides a direct test of this proposal by having participants give probability estimates of possible actions, using a method devised by Over, Hadjichristidis, Evans, Handley & Sloman (2007). Over et al. used conditional statements such as, "If the cost of petrol increases (p), then traffic congestion will improve (q)" and asked participants to estimate the probabilities of the four possible outcomes; p & q are true (TT), p is true & q is false (TF); p is false & q is true (FT),

and both p & q are false (FF). They found strong evidence in support of the conditional probability hypothesis with both a truth table task and a causal reasoning task. The conditional probability hypothesis states that a conditional statement, if p then q, is interpreted as the conditional probability of q, given p (P(q|p)) and can be calculated using probability estimates of truth and falsity, as in the following formula:

$$P(q|p) = P(TT) / [P(TT) + P(TF)]$$

Over et al. also found a small effect in support of the delta-p rule hypothesis, which considers conditional statements as the difference between the conditional probability and the probability of q, given not-p (P(q|p) – P(q|¬p)). Using their technique for estimating probabilities, the delta-p rule can be calculated using the following formula:

$$P(q|p) - P(q|¬p) = P(TT) / [P(TT) + P(TF)] - [P(FT) / [P(FT) + P(FF)]$$

Independent studies have supported the conditional probability hypothesis (Oberauer & Wilhelm, 2003; Ohm & Thompson, 2006). The research of Ohm & Thompson is particularly relevant as it applied the conditional probability hypothesis to inducements. They used conditional inducements and tips, similar to Newstead et al. (1997), and had participants judge the truthfulness of the conditional inducements and also the behavioural effectiveness (whether the inducement would bring about a change in behaviour). They predicted that conditional inducements and advice would be evaluated in terms of the conditional probability hypothesis (P(q|p)) but only to a degree. The pragmatic richness of inducements, they suggest, involves the representation of speech acts (Searle, 1969); the speaker utters q in order to increase or decrease the likelihood of the addressee performing p. This representation results in

instances of not-p being evoked. In fact, Ohm & Thompson appear to be suggesting that an inducement, e.g. "if you tidy your room, then you can play out", intuitively produces a response in the addressee to ask, "What happens if I don't tidy my room?" Such a view increases the relevance of not-p and can be evaluated using $P(q|\neg p)$ providing a measure of behavioural effectiveness.

Ohm & Thompson's predictions were supported in their findings; judgements of the truth of conditional inducements were predicted using the conditional probability hypothesis, while behavioural effectiveness correlated with the probability of q, given not-p. Interestingly, no correlation was observed between perceived truth and behavioural effectiveness, which Ohm & Thompson admit is counter-intuitive. They suggest that the majority of their statements were considered to be truthful and propose that a relationship may be found if they had included more conditionals with low truth values. Power of source is one means of manipulating both the perceived truth and perceived behavioural effectiveness of conditional inducements; high power should produce high truth and behavioural effectiveness, low power should be associated with low truth and behavioural effectiveness values. Indeed, many of the statements used by Ohm & Thompson involved power relations but they were not directly manipulated or made explicit.

The following experiment provides an opportunity to directly test the conditional probability hypothesis which thus far has been conjectured to be a possible explanation of pragmatic factors associated with deontic reasoning. Conditional and conjunctive phrased inducements comprising of threats, warnings, and two forms of promise were judged for likelihood of action. Previous studies have tended to classify

inducements into two groups: promises and threats forming one group and tips and warnings forming another (Fillenbaum, 1975, 1976; Newstead et al., 1997; Ohm & Thompson, 2006). Newstead et al. (1997) found much stronger effects with threats and promises on truth table and inference tasks. For this reason, a difference may be observed in the probability ratings of the inducements in this experiment. Inducements stated as conjunctive statements have not been used before in this type of experiment and, therefore, it will possible to test for any differences between conditional and conjunctive inducements.

The four scenario types (Family, Medical, Armed Forces, and Employment) are maintained in this experiment and although no significant differences between formal and informal scenarios were obtained previously, it was decided to test for significant effects because this experiment includes additional independent variables in which differences may emerge.

The credibility of the source of the message was also measured. It is proposed that manipulations of power of source will influence credibility ratings with individuals of high power expected to receive significantly higher credibility ratings compared to individuals of low power.

Finally, probability estimates will allow the calculation of conditional probability $(P(q|p))$ and the probability of q, given not-p, $(P(q|\neg p))$; a high power source will be associated with high $P(q|p)$ compared to a low source of power and a low source of power should produce high $P(q|\neg p)$ estimates compared to a high source of power.

From the estimates of $P(q|p)$ and $P(q|\neg p)$ the delta-p statistic can be derived which is predicted to be greater for high power.

Formally, the hypotheses can be stated as follows:

1. A high power source will be perceived as having significantly greater credibility compared to a low power source.

2. A significant difference will be observed in the credibility ratings of inducement types.

3. A significant difference will be observed in the credibility ratings for type of scenario.

4. Probability estimates of action will show significantly greater $P(q|p)$ values for a high power source compared to a low power source.

5. Probability estimates of action will show significantly greater $P(q|\neg p)$ values for a low power source compared to a high power source.

6. The delta-p statistic, $(P(q|p) - P(q|\neg p))$, will be significantly greater for a high power source compared to a low power source for both conditional and conjunctive inducements.

6.9 Experiment 8: Method

Participants

Participants comprised of 72 undergraduate students attending University College Northampton, with an age range of 18-60 years (median age 25 years). In one sitting 33 participants completed four experimental conditions containing *conditional* statements and in another sitting 39 participants completed four conditions containing *conjunctive* statements.

Design

A mixed experimental design was employed with four factors; social context, form of statement, type of inducement, and power of source. The social context factor comprised of four levels – family, health, armed forces and employment scenarios. The form of statement factor had two levels; conditional and conjunctive statements. Type of inducement consisted of four levels; threat, warning, promise (agential) and promise (external). The factor of power of source had two levels; high and low power. The independent variables (I.V.'s) for this experiment were scenario, statement form, type of inducement and power of source. The dependent variables (D.V.'s) were the credibility judgements and the probability estimates of possible outcomes of a message given its source.

Materials

The materials comprised of eight 10 page booklets, one for each experimental condition. The front page of each booklet introduced the task which stated,

Inside this booklet you will find a series of statements. Each statement involves one person making a threat, warning, or promise to another person. Given the statement and the person who makes it, you are asked to judge the credibility of the statement and to estimate the probable outcome.

Participants were asked to *complete <u>all</u> statements in the booklet.*

The following page provided an example of the two tasks. The example concerned a young girl and her homework and, although the wording varied according to the type

of inducement, the same example was used for each experimental condition. The example was introduced as follows:

The following statements involve one person [threatening, warning, promising] another. Given the statement and the person who makes it, you are asked to judge the credibility of the statement and to estimate the probable outcome.

A target statement then followed which varied according to the type of inducement. For the conjunctive conditions the statements read:

Threat: *Don't finish your homework and I will punish you*

Warning: *Don't finish your homework and your teacher will punish you*

Promise (agential): *Finish your homework and I will give you a treat*

Promise (external): *Finish your homework and your teacher will give you a treat*

The conditionals statements for each condition were as follows:

Threat: *If you don't finish your homework, then I will punish you*

Warning: *If you don't finish your homework, then your teacher may punish you*

Promise (agential): *If you finish your homework, then I will give you a treat*

Promise (external): *If you finish your homework, then your teacher will give you a treat*

One task for the participant was to judge the credibility of the statement, given the individual who makes it. The instruction read:

*Given that the [threat, warning, promise] is made by the young girl's **Mother**, how credible do you believe the statement to be?*

The credibility scale ranged from 1-10 where 1 was designated *NOT AT ALL CREDIBLE* and 10 *VERY CREDIBLE*. In the example the rating of 8 was circled and an explanation of the rating given:

If you judge the statement to be credible (that is, the mother means what she says) then you might choose to circle 8, as shown above

An example of the second task then followed. The instructions stated:

*Now consider the four possibilities that might occur following the statement. In each case estimate, in percentages, how probable you think each one is. **Make sure all four estimates add up to 100%***

The four possible outcomes were given with examples of percentage estimates. The wording of each example was consistent with the type of inducement and form of statement. The following is the example probability estimates for the warning inducement presented in the conjunctive form:

*The young girl <u>does not</u> finish her homework and the teacher <u>does</u> punish her
....35%....*

*The young girl <u>does not</u> finish her homework and the teacher <u>does not</u> punish her
....20%....*

*The young girl <u>does</u> finish her homework and the teacher <u>does</u> punish her
....5%....*
*The young girl <u>does</u> finish her homework and the teacher <u>does not</u> punish her
....40%....*

The participant was reminded that the probability estimates must total 100%. At this stage, before completing the experimental tasks, participants were asked if they clearly understood the task and the opportunity was given to ask any questions.

The next eight pages presented the materials for the participant to complete and comprised of four scenarios (Family, Medical, Armed Forces, and Employment) with individuals of low and high power embedded within each scenario. One of the eight booklets was completed by each participant comprising of one type of inducement (threat, warning, promise (agential) or promise (external) and one form of statement (conditional or conjunctive). The presentation of the scenarios within each condition was randomised to eliminate any order or learning effects.

The statements for each scenario used within each of the eight conditions is given below:

Conditional Threats

Family: *A boy is threatened: If you do not mow the lawn, then I will not give you your pocket money*

Medical: *A patient is threatened: If you don't stop drinking alcohol, then I will refuse to treat you*

Armed Forces: *An army volunteer is threatened: If you don't complete the training, then I will have you thrown out of the army*

Employment: *An employee is threatened: If you do not work the extra shift, then I will have you fired*

Conditional Warnings

Family: *A boy is warned: If you don't mow the lawn, then your mother may not give you your pocket money*

Medical: *A patient is warned: If you don't stop drinking alcohol, then the Health Service may refuse to treat you*

Armed Forces: *An army volunteer is warned: If you don't complete the training, then the General may throw you out of the army*

Employment: *An employee is warned: If you don't work the extra shift, then the company may fire you*

Conditional Promises (agential)

Family: *A boy is promised: If you mow the lawn, then I will give you your pocket money*

Medical: *A patient is promised: If you stop drinking alcohol, then I will treat you*

Armed Forces: *An army volunteer is promised: If you complete the training, then I will have you in my squad*

Employment: *An employee is promised: If you work the extra shift, then I will pay you double time*

Conditional Promises (external)

Family: *A boy is promised: If you mow the lawn, then your mother will give you your pocket money*

Medical: *A patient is promised: If you stop drinking alcohol, then the Health Service will treat you*

Armed Forces: *An army volunteer is promised: If you complete the training, then the General will have you in his squad*

Employment: *An employee is promised: If you work the extra shift, then the company will pay you double time*

Conjunctive Threats

Family: *A boy is threatened: Don't mow the lawn and I will not give you your pocket money*

Medical: *A patient is threatened: Stop drinking alcohol and I will treat you*

Armed Forces: *An army volunteer is threatened: Don't complete the training and I will have you thrown out of the army*

Employment: *An employee is threatened: Don't work the extra shift and I will have you fired*

Conjunctive Warnings

Family: *A boy is warned: Don't mow the lawn and your mother may not give you your pocket money*

Medical: *A patient is warned: Don't stop drinking alcohol and the Health Service may refuse to treat you*

Armed Forces: *An army volunteer is warned: Don't complete the training and the General may throw you out of the army*

Employment: *An employee is warned: Don't work the extra shift and the company may fire you*

Conditional Promises (agential)

Family: *A boy is promised: Mow the lawn and I will give you your pocket money*

Medical: *A patient is promised: Stop drinking alcohol and I will treat you*

Armed Forces: *An army volunteer is promised: Complete the training and I will have you in my squad*

Employment: *An employee is promised: Work the extra shift and I will pay you double time*

Conditional Promises (external)

Family: *A boy is promised: Mow the lawn and your mother will give you your pocket money*

Medical: *A patient is promised: Stop drinking alcohol and the Health Service will treat you*

Armed Forces: *An army volunteer is promised: Complete the training and the General will have you in his squad*

Employment: *An employee is promised: Work the extra shift and the company will pay you double time*

Each statement was followed by the ratings of credibility and probability estimates that the participant completed. See Appendix A8 (Experiment 8: Materials) for copies of the booklets used in this experiment.

Procedure

In this experiment, participants were presented with one of four inducements, in either the form of a conditional or conjunctive statement, and were required to rate the credibility of the statement given its source and to make probability estimates of possible outcomes.

The four inducement types; threat, warning, promise (agential) and promise (external) were operationally defined as in the previous experiment. Two forms of syntactical

inducement were employed, the conditional and conjunctive. Since promises cannot be expressed as a disjunctive, this syntactical form was omitted from this experiment. The social scenarios also remained the same as those used in the previous experiments reported in this chapter (family, armed forces, medical & employment), however, the levels of power were reduced to two (high and low) with the intermediary power source omitted.

Separate booklets were devised for conditional and conjunctive statements and for each of the inducement types. Thus, there were eight conditions employed: 2 forms of statement with 4 forms of inducement (see Materials section for details). Two judgements were required of the participant: first a rating of the credibility of the statement, given that it is made by a source of high or low power. The credibility rating was made on a ten-point scale, where 1= not at all credible and 10= very credible. The second task was to estimate, in percentages, the probabilities of the four possible outcomes, presented as truth functional possibilities (TT, TF, FT, FF).

To emphasise the truth or falsity of the antecedent and consequent, and to aid the reading of syntactically similar material, key terms were underlined. Participants were asked to estimate the likelihood of each outcome by assigning a percentage to each of the four possible outcomes; participants were reminded that the total estimates should equal 100%. All participants completed the tasks across all scenarios and for both cases of high and low power but for only one form of inducement. The ordering of scenario and power of source was randomised. No time limit was set for the completion of the two tasks.

6.10 Results

For raw data see Appendix B8 (Experiment 8: Raw Data). The experiment was completed satisfactorily by all participants and all data was retained for analysis.

Statistical analysis of the credibility data comprised of a three-way (2 x 2 x 4) mixed analysis of variance (ANOVA). The analysis was undertaken using SPSS (version 15). Mauchley's Test revealed that sphericity had been violated with the within participants factor of scenario and for the interaction of power x scenario, therefore, in these cases Greenhouse-Geisser values are reported. Significant effects were found for power of source ($F(1,62) = 269.69$, $p < 0.001$) and type of scenario ($F(3, 186) = 10.811$, $p < 0.001$). The interaction between power of source and type of scenario was also significant ($F(3,186) = 28.213$, $p < 0.001$). No significant effects were found for form of inducement ($F(7,62) = 0.987$, $p = 0.449$) or for the interactions between power of source x type of inducement ($F(7,62) = 0.927$, $p = 0.493$) and type of scenario x form of inducement ($F(21, 186) = 1.094$, $p = 0.358$). The interaction of power of source x type of scenario x form of inducement also showed no significant effect ($F(21,186) = 1.151$, $p = 0.299$).

In terms of the predictions for this experiment, hypotheses 1 and 3 are duly supported; a high power source is perceived as having greater credibility compared to a low power source and credibility ratings for the type of scenario did significantly differ. However, hypothesis 2 was not supported as no significant difference was observed in the credibility ratings of inducement form.

Conditional probability responses were analysed by comparing the probability estimates assigned to each of the four possible outcomes (TT, TF, FT, FF), for both conditional and conjunctive inducements. Conditional and conjunctive statements were analysed separately with data collapsed across type of inducement and form of scenario. Table 6.5 shows the mean probability estimates and standard deviations of the four possible outcomes, given the status of the speaker who makes the utterance for both conditional and conjunctive inducements.

Logical Cases (Possible Outcomes)				
	TT	TF	FT	FF
Conditional Inducements				
Low Power	28.27 (10.27)	21.82 (9.14)	10.43 (6.02)	39.78 (10.12)
High Power	39.20 (16.94)	12.60 (6.69)	7.10 (4.84)	41.26 (16.36)
Conjunctive Inducements				
Low Power	30.14 (8.28)	22.37 (11.42)	10.82 (6.09)	36.88 (11.23)
High Power	42.25 (15.38)	12.29 (6.89)	8.21 (4.88)	37.22 (15.03)

TT = true antecedent & true consequent; TF = true antecedent & false consequent; FT = false antecedent & true consequent; FF = false antecedent & false consequent

Table 6.5: Mean probability estimates and standard deviations (in parentheses) of the four possible outcomes for conditional and conjunctive statements made by a high or low power source

In estimating the probability of outcomes we are particularly interested in the situations in which the antecedent is true because this is were an authority figure can exert their influence. We would, therefore, expect to observe a greater probability for the TT case for a high power source than a low power source because a figure of

authority is more likely to make the consequent true. Conversely, the TF case is more likely to produce higher probability estimates from a low power source because they are less able to make the consequent true.

From the estimated probabilities the conditional probability was calculated using the formula: $P(q|p) = P(TT) / [P(TT) + P(TF)]$; the probability of q, given not-p was also calculated: $P(q|¬p) = [P(FT) / [P(FT) + P(FF)]$; and the delta-p statistic was derived using the formula:

$P(q|p) – P(q|¬p) = P(TT) / [P(TT) + P(TF)] – [P(FT) / [P(FT) + P(FF)]$

The mean and standard deviations of these calculations are shown in Table 6.6.

	$P(q|p)$	$P(q|¬p)$	$P(q|p) – P(q|¬p)$
Conditional Inducements			
Low Power	0.544 (0.202)	0.214 (0.129)	0.331
High Power	0.741 (0.160)	0.168 (0.126)	0.576
Conjunctive Inducements			
Low Power	0.581 (0.154)	0.237 (0.150)	0.353
High Power	0.757 (0.152)	0.227 (0.176)	0.533

Table 6.6: Mean values and standard deviations (in parentheses) of P(q|p), P(q|¬p), and P(q|p) – P(q|¬p) for conditional and conjunctive inducements made by a high or low power source

Correlated t-tests were used to test for a difference between the two sources of power. For conditional probability ratings ($P(p|q)$) significant results were obtained for both conditional ($t = 5.252$, $df = 32$, $p > 0.001$) and conjunctive inducements ($t = 5.254$, df

= 38, p < 0.001). Ratings of P(q|¬p) produced a significant difference for conditional

inducements (t = 2.322, df = 32, p < 0.05) but not for conjunctive inducements (t =

0.024, df = 38, p > 0.05). The delta-p calculations were found to produce significant

differences for both conditionals (t = 5.982, df = 32, p <0.001) and conjunctives (t =

4.181, df = 38, p < 0.001).

All significant results were in the predicted direction and supported the hypotheses,

with one exception. Hypothesis 4 was supported with significantly greater P(q|p)

values being attributed to a high power source compared to a low source of power and

this effect was observed for both conditional and conjunctive inducements. The

delta-p estimates were also significantly greater for a high power source compared to

a low power source, thus supporting hypothesis 6; a finding observed with both

conditional and conjunctive inducements. For hypothesis 5, probability estimates of

P(q|p) were significantly greater for a low authority source compared to a high

authority source for conditional inducements only; conjunctive inducements produced

no significant difference.

For SPSS output, see Appendix D8 (Experiment 8: SPSS Output).

6.11 Discussion

The three experiments reported in this chapter demonstrate the influence of power of

source as an additional pragmatic factor when reasoning with deontic statements.

Experiment 6 showed that the authority of an individual making a conditional

statement influences the perceived likelihood of action, across low, medium and high

power sources. This finding was repeated across a range of scenarios, however, the

predicted difference between formal and informal social contexts did not hold; it appears that a powerful source will wield their influence whatever the formality of the social situation.

The findings of Experiment 7 demonstrate just how general power of source is in deontic contexts. The experiment extended the syntactic form, to conjunctive and disjunctive statements and extended the type of inducement by incorporating promises, threats and warnings. Significant trends were found in the predicted direction; the likelihood of an inducement being made true increased with perceived authority. Only the Armed Forces scenario showed anomalies in predicted trends, occurring with the inducements of warning and threat.

Experiment 8 tested the conditional probability hypothesis that participants are considering the probability of an individual making the consequent (q) of a statement true, given its antecedent (p) with power of source as an intervening variable. Inducements made by an individual of high social rank were rated as more credible than if made by an individual of low social rank. No statistical difference was observed between statements presented as conditional or conjunctive inducements, suggesting both syntactic forms evoke deontic reasoning.

Estimates of the probability of outcomes provided a method for directly assessing the conditional probability hypothesis. From the probability estimates, $P(q|p)$ and $P(q|\neg p)$ were derived with the former showing higher mean values for a high power source and the latter higher mean values for a low power source, as predicted. Significant findings were obtained for both conditional and conjunctive inducements with one

exception; no difference was observed in the P(q|p) estimates between high and low sources of power for conjunctive inducements. The delta-p value was also calculated and predicted to be significantly higher for high ranking power compared to low power. A significant effect was observed with both conditional and conjunctive inducements.

Overall, this provides good support for the conditional probability hypothesis, with the only anomaly being the P(q|¬p) measure for conjunctive inducements. Over et al. (2007) found the strongest predictor of performance on their tasks to be the probability of q, given p (P(q|p)) with the delta-p rule (which includes P(q|¬p) in its calculation) only having a marginal influence. According to the conditional probability hypothesis statements made by low power sources should increase the number of generated counter examples which will be manifested in the P(q|¬p) measure. This was observed (see Table 6.6), however, the high power source for conjunctive inducements also produced a relatively high P(q|¬p) value resulting in a non-significant effect.

There may be an explanation for the non-significant result with the conjunctive probability estimates of q given not-p. Ohm & Thompson (2006) found high truth values assigned to the majority of inducements they used in their research; Experiment 8 also found relatively high conditional probability estimates for low power. It, therefore, appears that listeners will assume a message is true, as suggested by Grice's maxim of *quality* and that may include a message from a low power source. Indeed, low ranking individuals may provide felicitous statements,

particularly when expressed as a tip or warning. Only when the message or source of the message is questioned does the process of generating counter examples begin.

Of course there may be occasions when an individual of low rank has sufficient authority to make a true assertion, for example when stating a warning or giving advice. This would raise the question of the felicity of the utterance and may suggest why warnings are often classed as a separate class of inducement from promises and threats. This will require further experimentation to isolate any additional pragmatic factors that may be in operation.

The findings reported here can be contrasted with those of Ray et al. (1989) who failed to capture any effects of power. As suggested previously (see section 6.1) Ray and colleagues were not directly investigating power relations and the authority figures used in their scenarios are questionable. The experiments reported in this chapter show how the influence of power can be manipulated to produce significant effects. Similarly, with previous studies of inducements (Newstead et al., 1997; Ohm & Thompson, 2006) some of their findings may be associated with the power relations involved in the social scenarios employed.

The social scenario and the roles of individuals within it bring to the fore a number of other pragmatic issues. For instance the anomalies observed in the ratings of the armed forces scenario in Experiments 6 and 7 may be explained in terms of participants being unable to distinguish between the ranks of Sergeant and Corporal, either because both ranks are considered to hold significant power given the strict hierarchical structure of the armed forces, or else ignorance of which is the higher

ranking officer. It could, therefore, be argued that the Armed Forces scenario did not sufficiently capture the power differences in rank with the statements used. The Armed Forces scenario may be at the opposite extreme to the contexts used primarily by Ray et al. (1989); within the Armed Forces, all offices of rank have the ability to make a statement true.

In sum, the experiments reported in this chapter have shown that power of source is a further pragmatic factor that has broad generality. A range of social contexts encompass power relations and much of our discourse involves inducements. Inducements themselves can be expressed as conditional, conjunctive and disjunctive statements which appear to be understood equally well. The suppositional reasoning theory of Evans and colleagues (Evans, Handley & Over, 2003; Evans, Over & Handley, 2003; Evans & Over, 2004) offers a better explanation of the findings as it emphasises the probabilistic nature of reasoning. This theory would suggest that System 1 is active when reasoning with inducements and the pragmatics of power. Participants are determining the probability of the consequent of a statement being made true, and this is most likely when the source of the statement is of high social rank. The findings, therefore, support a growing body of research suggesting reasoning, including deontic reasoning, is probabilistic.

Chapter 7

Judgement Revision with Power of Source and Inducements

7.1 Rationale

Power of source has been shown to be a pragmatic factor of deontic reasoning. The experiments of the previous chapter demonstrated the generality of an authority source influencing the acceptance of conditional statements, inducements, and conjunctive and disjunctive assertions. An utterance made by a high power source is perceived as having more credibility than an utterance from a low source of power. Further, the probability of outcomes is perceived to be controlled by a high ranking individual to a greater extent than a low status individual, as shown with measures of conditional probability $(P(q|p))$. There is also evidence to suggest that when considering a statement by a low power source the number of counter-examples increases, particularly with conditional inducements, as shown in the measure of the probability of q, given not-p $(P(q|\neg p))$. This chapter develops the pragmatic factor of power of source by embedding it in a judgement revision task.

The proposal that reasoning and decision-making are invariably linked has gained gradual credence. Evans, Over & Manktelow (1993) suggested both reasoning and decision-making are concerned with rationality. Studies with thematic versions of the selection task involved features associated with decision-making, such as probability (Kirby, 1994; Manktelow, Sutherland & Over, 1995) and utility (Manktelow & Over, 1991). The realisation that a common feature of many reasoning tasks was the involvement of deontic reasoning processes (Cheng & Holyoak, 1985) also increased the credibility of the argument that reasoning and decision-making were associated.

Deontic reasoning involves permissions and obligations which are not truth functional like indicative logical reasoning tasks, and allow the inclusion of probabilities. Thus, Oaksford & Chater (1994) and Manktelow & Over (1995) called for a decision-theoretic approach to deontic reasoning.

Theoretically, developments have moved in the direction of a probability theory of reasoning. Oaksford & Chater (1994) proposed a Bayesian theory to explain the findings with the selection task, called *Information Gain* theory. They have extended their probabilistic theory to explain a number of related areas, such as conditional reasoning (Oaksford, Chater & Larkin, 2000) and syllogistic reasoning (Chater & Oaksford, 1999).

Inspired by the work of Oaksford and Chater, Evans and colleagues have developed a separate probabilistic theory. The suppositional theory (Evans, Handley & Over, 2003; Evans, Over & Handley, 2003; Evans & Over, 2004) proposes that utterances are considered in terms of the probability of the antecedent (p) and the consequent (q), in particular the conditional probability i.e. the probability of q, given p (P(q|p)). Support for the theory is steadily growing and has been used to explain a number of phenomena, including indicative conditionals (Oberauer & Wilhelm, 2003; Over, Hadjichristidis, Evans, Handley & Over, 2007), inducements (Ohm & Thompson, 2006) and causal and diagnostic conditionals (Evans, Handley, Hadjichristidis, Thompson, Over & Bennett, 2007). Experiment 8, reported in the previous chapter, provided a test of the theory in relation to power of source with positive results.

The empirical investigation and theoretical developments in decision-making and judgement revision has produced two distinct approaches; heuristics and linear models (Hogarth & Karelaia, 2007). The heuristics approach is based on the view that human cognitive capacity is limited and therefore simple heuristic are employed when making decisions (Gigerenzer, Todd, & the ABC research Group, 1999; Kahneman, Slovic & Tversky, 1982). The approach of Kahneman, Tversky and their collaborators has been to show that human reasoning and decision-making does not meet the standards of normative theories (Tversky & Kahneman, 1973, 1974). Gigerenzer's approach has been to obtain evidence for proposed heuristics, such as *take-the-best* (TTB) (Gigerenzer & Goldstein, 1996). This idea has intuitive appeal, however, a difficulty for the approach is predicting under what environmental conditions heuristics will be employed.

The linear models approach to judgement and decision-making has been influenced by the work of Brunswik (1952) and assumes the decision-maker can hold all information. This information, or environment cues, are combined and weighted to produce a *rational* judgement or decision. Several linear models have been proposed (Dawes, 1979; Dawes & Corrigan, 1974; Einhorn & Hogarth, 1975; Keeney & Raiffa, 1976). The advantage of this approach is that the inclusion of environmental cues allows for the weighting of information in terms of content and context. A disadvantage is that linear models are based on algebraic calculations and, therefore, assume a degree of System 2 thinking. However, the two approaches, heuristics and linear models, are not mutually exclusive as many theorists today believe it to be perfectly reasonable to assume that judgement and decision-making can involve both processes (Hogarth & Karelaia, 2007).

An additional manipulation is to be incorporated in to the experiments of this chapter which relate to a further proposal for explaining the range of mediating pragmatic factors influencing reasoning in everyday contexts. Fairley & Manktelow (2004) and Manktelow & Fairley (2000) suggest everyday reasoning contexts invoke superordinate principles (or Super P's). Such principles are drawn from one's personal knowledge and experience and capture the necessary and sufficient conditions for a particular outcome. An example taken from Manktelow & Fairley (2000, Experiment 1) will illustrates their idea.

A familiar conditional statement was set in syllogistic form, "If you study hard, you do well in exams". Knowing that a student has studied hard, one would make the logical modus ponens conclusion that the student will do well in exams. However, a number of disabling conditions may influence the sufficiency of the deduction drawn (for instance, that the student arrived late for the exam or is very nervous). Similarly, with the denial of the antecedent (DA) argument form the statement, "If you study hard, you do well in exams" and if then informed, "You do not study hard", then a common but logically fallacious conclusion may be drawn, "You do not do well in exams". With this form of argument, alternative causes were shown to influence the necessity of the drawn deduction (for instance, if told that the student cheated or was intelligent). A control condition was also introduced in which irrelevant factors were stated (for instance, the student wore blue socks or was friendly).

In the conditional arguments above Manktelow and Fairley suggest "doing well in exams" may be interpreted as a Super P of the form, "ways of passing an exam". Studying hard may be a pre-requisite for passing exams, however, one's knowledge of

such contexts may provide a set of circumstances (alternative causes) whereby the outcome is made true. Similarly, past knowledge and experience will make one aware of possible disabling conditions, whatever preparations may have been undertaken, resulting in failure.

In a further experiment (experiment 3) Manktelow and Fairley used Manktelow & Over's (1991) well known conditional statement, "If you tidy your room, then you may go out to play". Alternative antecedents were introduced in the form of, "tidying the living room" or "washing the dishes" which resulted in the participants believing that the conditions had been satisfied to make the conclusion true, that is, the child could go out to play. Disabling conditions such as, "watching TV" or "phoning a friend" were considered as negations of the antecedent and, therefore, the outcome of going out to play was not made true. Irrelevant factors, such as "completed homework" or "made a cake" made no direct impact on the conclusion drawn. Thus, a general Super P appears to be in operation in this context, that of "helping around the house". In sum, alternative causes (antecedents) are those factors that satisfy Super P, while disabling conditions (antecedents) are those that prevent Super P from being satisfied.

Manktelow & Fairley (2000) modified their sufficiency / necessity theory when explaining superordinate principles by couching their explanation in probabilistic terms. They suggest that alternative causes increase the probability that an event will follow in the absence of the cause and, therefore, undermine necessity while disabling conditions and additional requirements both decrease the probability that an event will follow the cause, that is, they undermine sufficiency. The experiments reported in

Manktelow & Fairley (2000) provide strong evidence for the existence of superordinate principles in both the causal and deontic domains and their revised theory is more consistent with current thinking and the approach taken within this thesis.

In the experiments of Manktelow and Fairley, Super P's were inferred from subordinate statements; they did not use explicit superordinate statements. In the following experiment superordinate conditional statements will be compared with typical, non-superordinate, if…then statements. This will allow for a better test of the proposal for superordinate principles within the domain of deontic reasoning. Power of source seems a suitable context in which to expect superordinate principles to exert an influence; an individual with authority should be more influential with both general and superordinate conditional statements.

Judgement revision was demonstrated within the context of motoring offences in Experiment 5 of this thesis (see Chapter 5). Judgements were revised in predictable directions when information concerning the nature of offence, scale of violation, and accompanying mitigating or aggravating circumstances was presented in series (i.e. step-by-step, cf. Hogarth & Einhorn, 1992). In particular, significantly revised judgements were most likely to be observed when a *contrast effect* was present, for example, a minor scale of violation followed by aggravating circumstances (i.e. 35 mph – disqualified driver). It is predicted that in the current experiment similar effects will be observed within the context of inducements; in the scenarios used scale of violation and power of source will provide the subsequent information for revised judgements.

In sum, the following experiment involves judgement revision within the context of three types of inducement (threat, warning and promise). Two scenarios are utilised; inducements uttered within a family setting to a teenage son and inducements made to employees within a company. A further manipulation is the form of statement spoken to the teenage son or employee in order to measure perceived differences between a general superordinate statement and a more specific non-superordinate utterance. A superordinate statement refers to *misbehaviour* or *well behaved* within the family context and *misconduct* and *excellent behaviour* within the company scenario. The non-superordinate statements relate to lateness within the family setting and time-keeping within the context of a company. Three judgements of likelihood are measured; the initial rating, given the inducement and the behaviour of the teenager or employee; a second judgement, given the scale of violation (for threat or warning) or rate of compliance (for promise); a third rating, given the authority of the individual making the initial statement. Scale of violation was categorised as major or minor and power of source as high or low.

It was predicted that if utterances comprise of a hierarchical structure with superordinate principles encompassing many specific statements a difference would be expected in the perception of such statements which would be reflected in the likelihood ratings of action. A difference is therefore predicted between superordinate and non-superordinate statements. A difference is also expected between a major and minor violation (or compliance) with likelihood ratings of action increasing for major transgressions. Power of source is predicted to result in a high power source producing greater likelihood judgements of action than a low power source. Formally, the hypotheses can be stated as follows:

1. A significant difference in the likelihood ratings of action will be observed between superordinate conditional statements and non-superordinate statements.

2. A major transgression (compliance) will produce greater likelihood judgements of action compared to a minor transgression (compliance).

3. A high power of source will produce greater likelihood ratings of action compared to a low power source.

7.2 Method

Participants

Participants comprised of 457 post-16 students attending Stourbridge College of Further Education and undergraduate students attending the University of Wolverhampton. A further 18 participants were not used in the final analysis because they had not completed all judgement revision tasks. Of those completing the required tasks 154 were males (age range 18 - 45 years) and 303 females (age range 16 - 46 years). All participants were unpaid volunteers and allocated to one of 48 conditions.

Materials

The materials comprised of A5 size booklets containing one of 48 conditions. The first page provided instructions for the task and an example. The instructions stated:

> *For this task, you will be presented with a statement that is spoken by one individual to another. Given the statement you must judge the likelihood of action. Further details will be given on subsequent pages and you will have the option of revising your decision.*

An example then followed:

> *For example, given the following statement spoken to a girl:*
> ***"Finish your homework and I will take you to the cinema"***

Given the girl enjoys the cinema, how likely is it that she will finish her homework?

Very Likely   *[ ]*
Likely   *[✓]*
Unsure   *[ ]*
Unlikely   *[ ]*
Very Unlikely   *[ ]*

If you believe the girl is likely to finish her homework, you would tick the Likely box, as shown above.

Please turn each page in turn and make your judgement before moving to the next page.

The details of each inducement was presented over three pages with the inducement given first, followed by the scale of violation or compliance and finally the authority of the individual making the statement, either of high or low power.

An example of the materials used for the judgement revision task is given below (this example is for the non-superordinate condition with family scenario and threat induction):

A family is concerned about their teenage son who is repeatedly arriving home late in the evening. In an attempt to improve his time-keeping a threat is made to ground the teenager if he arrives home after 11pm during the following month.

The teenage son is given the following threat: ***"If you arrive home after 11pm during   the following month, then I will ground you"***

The teenage arrives home after 11pm during the following month

How likely is it that the teenager will be grounded?

Very likely   *[ ]*
Likely   *[ ]*
Unsure   *[ ]*
Unlikely   *[ ]*
Very Unlikely   *[ ]*

The above was Rating 0 (R0), prior to any revision of judgement. The first revised judgement was labelled R1 (Rating1) and detailed the scale of violation. For the example above a minor violation was for the teenager to arrive home 10 minutes late; a major violation, arriving home 3 hours late. Rating (R2), the second revised judgement, involved information about the source of the initial inducement, either an individual of high or low authority (for the example above, father was used as high power and younger brother as low power). Each judgement involved assessing the likelihood of action on a 5-point Likert scale, as shown above. For the family threat example, participants were judging the likelihood of the teenager being grounded.

For the full set of materials see Appendix A9 (Experiment 9: Materials).

Design

A between-participants design was employed with three factors having two levels each; principle (superordinate and non-superordinate), scale of violation/compliance (major and minor), and power of source (high and low). The factors of inducement and scenario were collapsed following the non-significant findings of the previous experiments. The independent variables for this experiment were, therefore, principle, scale of violation/compliance and power of source. Three dependent variables were measured, the initial judgement of likelihood given the inducement (R0), the revised judgement of likelihood of action given scale of violation/compliance (R1), and the second revised judgement given the power of the message source.

Procedure

Participants were tested in large groups, ranging in size from 20 – 100, and randomly assigned to one of 48 conditions. The conditions of the experiment comprised principle (superordionate / non-superordinate), inducement (threat, warning, promise), scenario (family, company), scale of violation/compliance (major / minor) and power of source (high / low); hence, 2 x 3 x 2 x 2 x 2 = 48.

The superordinate condition for the family threat and warning scenarios referred to *misbehaviour*, while the promise condition referred to *behaving well*. The non-superordinate family conditions referred to *arriving home late*. Thus, the superordinate conditions used a general reference of behaviour, the non-superordinate conditions used a specific behaviour reference, lateness. Similarly for the company scenario, superordinate threats and warnings referred to *misconduct*, while the promise refers to *excellent behaviour*; non-superordinate threats and promises involve *arriving late for work* and the promise refers to *excellent time-keeping*.

Three inducement types were used, threat, warning and promise (agential). The promise made by an external agent was not used for this study. Each inducement was stated as a conditional, if…then, statement with disjunctives and conjunctives being omitted.

The scenarios of Family and Company were employed with company being a variation on the employment scenario used in previous experiments. The scenarios of Armed Forces and Medical were not employed.

Scale of violation comprised of major and minor. For non-superordinate threats and warnings for the family scenario, a major violation involved a teenager arriving home 3 hours late; a minor offence arriving home 10 minutes late. Similarly for the company scenario, the employee arriving 10 minutes late for work constituted a minor violation and 3 hours late a major violation. Promises were stated in terms of compliance with the family scenario giving the inducement:

> *If you arrive home before 11pm for the following month, then I will give you an increase in your allowance*

A major compliance to this promise was that the teenager arrives home *every* evening for a month; a minor compliance was for the teenager to arrive home before 11pm *most* evenings. The company scenario involved a similar inducement concerning time-keeping with the employee described as arriving on time *every day* or *most days*, representing a major and minor compliance, respectively.

For the superordinate conditions the family scenario referred to *misbehaviour* when uttering a threat or warning, with the major violation being the teenager taking £20 without asking and a minor violation the teenager taking £1 without asking. The superordinate promise for the family context was expressed in terms of the teenager being *well behaved* for the coming month with a major compliance being an 100% improvement in the teenager's behaviour, a minor compliance a 50% improvement. The company threats and warnings, for the superordinate conditions, related to misconduct with the major violation being that the employee had stolen money and a minor offence the employee had taken an envelope belonging to the company. The

superordinate promise in the company context referred to *excellent behavioural conduct*, with high compliance being the employee won the Employee of the Month award and minor compliance the employee was nominated for the award.

The power of source remained constant across all conditions. For the family scenarios Father was used as the high power source and Younger Brother the low power source. The company scenario had the Managing Director as the high power source and Supervisor as the low source of power.

Each participant was given one of 48 booklets to complete. Each booklet was of A5 size and comprised four loose-leaf pages with text printed on one side only. Participants were asked to read the first page of the booklet which provided standardised instructions and an example to illustrate the forthcoming task (the same example was given in each of the 48 booklets). As the task involved judgement revision participants were asked to complete the tasks in strict order and not to "look ahead". The opportunity was given for participants to ask questions, if necessary.

Three judgements of action were made by each participant on a 5-point Likert scale., ranging from Very Likely to Very Unlikely. The responses were scored as shown below:

Very likely   [5]
Likely   [4]
Unsure   [3]
Unlikely   [2]
Very unlikely   [1]

No time limit was set for the completion of this task.

7.3 Results

The data from 18 participants was not used for analysis because they did not complete all three judgements. For raw data see Appendix B9 (Experiment 9: Raw Data).

Overall mean likelihood ratings are shown in Tables 7.1 and 7.2 for the non-superordinate and superordinate inducements, respectively (for full set of mean scores see Appendix C9 (Experiment 9: Descriptive Statistics)).

Non-superordinate Inducements

Family Scenario

Inducement	Major Violation/ Compliance	Minor Violation/ Compliance	Power of Source High	Low
Threat 3.7	4.4	2.6	4.1	1.5
Warning 3.3	3.9	2.5	3.9	2.1
Promise 3.9	4	3.3	3.5	2.8

Company Scenario

Threat 3.5	3.8	2.9	3.9	2.3
Warning 3.1	3.7	3.1	3.6	2.4
Promise 4.3	4.0	2.9	3.8	3.4

Table 7.1: Overall mean likelihood ratings for the non-superordinate inducements for family and company scenarios. Violations are associated with threats and warnings, compliance with promises.

Superordinate Inducements

Family Scenario

	Major Violation/ Compliance	Minor Violation/ Compliance	Power of Source High	Low
Threat				
3.4	4.1	2.9	3.7	1.8
Warning				
3.1	4.1	3.0	4.1	1.4
Promise				
4.1	4.3	3.4	4.8	2.1
Company Scenario Threat				
4.2	5.0	2.6	3.9	3.1
Warning				
4.0	4.6	2.7	3.9	2.7
Promise				
3.5	3.9	3.7	3.9	3.0

Table 7.2: Overall mean likelihood ratings for the superordinate inducements for family and company scenarios. Violations are associated with threats and warnings, compliance with promises.

The overall mean ratings for major violation / compliance are greater than the judgements for a minor violation or compliance, as predicted. Further, the likelihood ratings of action for a high power source are greater than a low source of power. These findings are consistent across type of inducement, scenario and super- and non-superordinate principles. The initial likelihood ratings are also very similar for superordinate and non-superordinate conditions; ranging between 3.1 – 4.3 for the non-superordinate inducements, and 3.1 – 4.2 for inducements presented in the superordinate form.

The data was analysed statistically using a three-way (2 x 2 x 2) between-participants analysis of variance (ANOVA). The analysis was undertaken using SPSS version 15. Three dependent variables were measured corresponding to the initial judgement (R0) and two revised judgements (R1 and R2). The first rating (R0) provides a comparison of inducements stated in superordinate and non-superordinate form; the second rating (R2) compares major violation / compliance with minor violation / compliance; and the final judgement gives a comparison of high and low power.

For the initial judgement (Super P v non-Super P) no significant effect was observed $(F(1,449) = 1.422, p = 0.234)$.

For the first revised judgement (R1: major v minor violation/compliance) a significant effect was found $(F(1,449) = 151.849, p < 0.001)$ with no interaction effect observed between scale of violation v principle $(F(1,449) = 0.725, p = 0.395)$.

For the second revised judgement (R2: high v low power source) a significant effect was observed $(F(1,449) = 177.589, p < 0.001)$. However, no significant interaction effects were observed between power of source v principle $(F(1,449) = 2.745, p = 0.098)$, power of source v scale of violation $(F(1,449) = 1.986, p = 0.159)$ or principle v power of source v scale of violation $(F(1,449) = 2.551, p = 0.111)$.

7.4 Discussion

The experiment in this chapter involved judgement revision of conditional inducements presented in superordinate or non-superordinate form. Manipulations of scale of violation / compliance and power of source provided embedded pragmatic factors that had previously been shown in previous chapters to influence deontic reasoning. A novel factor of the research was the inclusion of a superordinate level of inducement.

The findings provide support for two of the three hypotheses. There was no significant effect for principle (superordinate v non-superordinate) conditions. However, significant effects were observed for the factors of scale of violation / compliance and power of source.

The superordinate phrase provided a generic inducement compared to a more specific inducement with the non-superordinate conditions. It appears that either form produces similar likelihood ratings, however, the finding contrasts with the hypothesis which predicted a difference in ratings of likelihood. It may be that Manktelow and Fairley's (2000) claim that superordinate principles are invoked in many reasoning contexts means that the generic inducement form will not be easily distinguished from the more specific inducement or when the superordinate is stated explicitly. Further investigation may be required to disentangle the reasoning processes occurring with superordinate statements from indicative forms.

Scale of violation and power of source were observed to influence ones judgement which is consistent with the findings reported in Chapter 6. In the previous

experiments, however, judgements were made after all information was given (i.e. end-of series). In this experiment the influence of both scale of violation and power of source was observed in a judgement revision task (i.e. step-by-step). Thus it appears that these features are reliable pragmatic factors of both deontic reasoning and judgement revision.

The findings provide good evidence to support the view that reasoning and decision-making are associated and involve many of the same processes, as suggested by a number of theorists (Evans, Over & Manktelow, 1993; Manktelow & Over, 1995; Oaksford & Chater, 1994). If comparable findings can be obtained between reasoning tasks and judgement revision tasks, as reported here, it would suggest that similar cognitive processes are involved in both activities, at least for deontic contexts.

Although the findings support the view for a decision-theoretic approach to deontic reasoning (Manktelow & Over, 1995), they do not directly test any theory of human reasoning or decision-making. Therefore the experimental results may be explained by a number of explanations. Schema theories (Cheng & Holyoak, 1985; Cosmides, 1989) do not incorporate judgement revision into their models, however, revised schema may be plausible as subsequent information is received. Heuristic models of decision-making (Gigerenzer et al., 1999; Kahneman et al., 1982) could not currently explain the findings of this experiment without devising a further heuristic. Linear decision-making models, such as Brunswik's (1952) and derived models, may offer a more appropriate explanation as they incorporate environmental variants, which could include pragmatic effects. The linear models may also share more features with

probabilistic accounts of reasoning (Oaksford & Chater, 1994; Evans, Handley & Over, 2003) as they are both algebraic and can incorporate measures of probability and utility. The design of the experiment in this chapter would have to be adapted to test a particular theory directly, however, this is feasible.

A number of features of the experiment could be improved. The range of scenarios were rather limited, however, the chosen contexts had produced reliable findings in previous experiments. Inducements produce different outcomes with promises offering a reward, while threats and warnings propose a sanction. The inclusion of a promise also meant that the judgement task involved rating the likelihood of an individual complying with an inducement, rather than violating it as occurs with threats and warnings. The classification of major and minor violations have been demonstrated in earlier experiments using motoring transgressions, however, rating major and minor compliance may be quite a different task from considering scaled violations.

In sum, the findings showed no difference in the ratings of superordinate and non-superordinate statements which has implications for Manktelow & Fairley's (2000) proposal for a hierarchical structure to spoken inducements. Significant effects were found for scale of violation and power of source, suggesting these pragmatic factors are transferable from reasoning to judgement revision tasks. The findings give support to the view that reasoning and decision-making involve similar cognitive processes but do not provide specific support for any particular theory of either reasoning or decision-making.

Chapter 8

General Discussion

8.1 Overview of Chapter Content

This chapter will begin with a summary of the experimental findings of the experiments reported in the thesis. The implications of the findings will then be discussed in relation to current issues and thinking within the reasoning and decision-making fields. A theoretical discussion will then consider current deontic reasoning theories as they apply to the findings of this research and it will be postulated that a probabilistic theory best explains the reported results. Finally, proposals for further research will be made in order to extend and develop the research programme of pragmatic factors of deontic reasoning.

8.2 Summary of Experimental Findings

The first pragmatic factor to be investigated was scale of violation within a *working-rule* scenario (Experiment 1, Chapter 3). It was observed that participants rate the breaking of a rule as more serious during the day of the violation (Day 1) rather than the extent to which the rule is violated on a subsequent day (Day 2). Mean fines increased for Day 1 as the scale of violation increased over 1, 4 & 8 hours. However, the mean fines for Day 2 showed little difference as scale of violation increased.

The findings of Experiment 1 involved the use of an inference task. The aim of the next experiment (Experiment 2, Chapter 3) was to attempt to repeat the results using a deductive reasoning task, specifically a Large Array Selection Task (LAST) (Manktelow, Sutherland & Over, 1995). However, the findings in Experiment 1 were

found not to transfer to the LAST and it was suggested that the detection of violators in such a task was relatively easy and did not provide discriminations of judgement.

Experiment 3 (Chapter 4) considered a further pragmatic factor, the circumstances associated with violations. The contexts used for this experiment was motoring transgressions; speeding and drink-driving. Significant effects were observed across three classes of circumstances; mitigating, neutral and aggravating. A significant difference in *seriousness* ratings was also observed between the two offences, with drink-driving considered the more serious.

Having established significant findings for scale of violation (Experiment 1) and circumstances (Experiment 3), the two pragmatic factors were combined in Experiment 4 (Chapter 4). Using the motoring transgressions of speeding and drink-driving, scale of violation was incorporated with the use of a major and minor offence with additional aggravating and mitigating circumstances. Participants judged whether a motorist should be fined, using a list of modal terms (e.g. must, should, must not, should not etc.). Significant effects were found for circumstances, type of motoring offence, and scale of violation providing further support for considering these effects as general pragmatic factors of deontic reasoning.

The subsequent experiment investigated the relationship between reasoning and decision-making with material presented in series (Experiment 5, Chapter 5). This method of presentation affords the opportunity for participants to revise their judgements. In this task a fine was determined on a scale from £0 (no fine) to £100 (max. fine) with the opportunity for two revised judgements, given the scale of the

violation and the circumstances. Significant differences were found for the initial judgement involving the transgression (speeding v drink-driving), for the first revised judgement when scale of violation was introduced (major v minor) and for the second revised judgement with the introduction of circumstances (aggravating v mitigating). Post hoc t-tests revealed significant differences between revised judgements, particularly when a contrast effect was apparent, for instance a minor offence followed by aggravating circumstances. In general, judgement revision followed predicted directions, given the results of the previous experiments.

A further pragmatic factor, power of source, was introduced in the next series of experiments. Experiment 6 (Chapter 6) used conditional statements within an inferential reasoning task and it was found that an individual of high social rank was judged to be more likely to make a statement true compared to an individual of low social rank and this finding was consistent across both formal and informal scenarios. Significant trends from low-high power were observed for all scenarios, except the armed forces context, which produced an anomalous result between medium and high power.

The influence of power was then extended to three types of inducement (threat, warning and promise) and to conjunctive and disjunctive statements (Experiment 7, Chapter 6). The findings with inducements were similar to those obtained with conditional statements; individuals of high power were perceived as having a greater probability of making the consequent true given the antecedent and this was also found for conjunctive and disjunctive statements. Significant trend tests were

observed for each inducement (threat, warning, promise agential and promise external) across scenario (family, medical, armed forces, and employment).

Experiment 8 (Chapter 6) provided a theoretical-basis to the research with power of source by directly testing the conditional probability hypothesis (Evans, Handley & Over, 2003; Evans & Over, 2004). Participants provided ratings of credibility given the authority of the message source and probability estimates of possible outcomes. Power of source was again found to produce a significant effect with a high power source rated as more credible than a low power source when uttering both conditional and conjunctive inducements. Probability estimates were used to calculate conditional probability ratings (P(q|p)) which were found to be significant for both conditional and conjunctive inducements, indicating that a high power source was more likely to make q (the consequent) true, given p (the antecedent) compared to a low ranking source. Further, calculations of (P(q|¬p)) indicated that counter-examples (not-p's) were significantly more likely to occur when considering a conditional inducement spoken by a low power source but not when considering conjunctive statements. The delta-p statistic, (P(q|p) – P(q|¬p)), would predict a greater value for a high power source compared to a low power source, and this was observed for both conditional and conjunctive inducements.

The final experiment (Experiment 9, Chapter 7) involved a judgement revision task incorporating inducements, scale of violation with threats and warnings or compliance with a promise, and power of source. A judgement of the likelihood of action was made for the initial inducement (threat, warning, promise agential) and for two revised judgements – scale of violation/compliance (major or minor) and power of

source (high or low). A significant difference was observed for the revised judgements; a major violation/compliance producing a greater likelihood of action than a minor violation/compliance and a high power source producing a greater rating than a low source of power. A further manipulation comparing superordinate statements (Fairley & Manktelow, 2004) with non-superordinate inducements did not produce a significant effect.

8.3 Implications of Experimental Findings

The findings from the experiments conducted for this thesis have a number of implications for the thinking and reasoning literature. In particular we shall consider how pragmatic factors develop earlier findings of content effects and thematic versions of the Wason selection task. Pragmatic effects also extend the range of factors found to modify deontic reasoning. The context in which pragmatics operate must also be stated; pragmatics involve familiar, real-world situations and involve practical, System 1 thinking. Pragmatics can also be applied to judgement revision tasks bringing the fields of reasoning and decision-making closer together. Finally, some consideration will be given to the unpredicted and anomalous findings within the reported experiments.

Content and context effects with syllogistic and other forms of reasoning were documented from the earliest studies (Wilkins, 1928) and facilitation with thematic versions of the Wason selection task (Wason, 1968) suggested that reasoning cannot be devoid of everyday knowledge. Considering pragmatic factors is one way of investigating and explaining the use of world knowledge in reasoning and decision-making. Pragmatics within the study of language is well developed and it is suggested

here that such factors can be applied to reasoning and decision-making.

The seminal paper of Cheng & Holyoak (1985) provided a major development with the realisation that thematic versions of the selection task produced facilitation when they involved deontic reasoning. Deontic reasoning is concerned with permissions and obligations, a feature of the majority of the thematic selection tasks. Aristotle made the distinction between theoretical and practical reasoning with deontic reasoning being a special type of the latter. Practical reasoning is associated with one's goals, including utilities and involves inferential reasoning which is not truth functional. Further investigations into deontic reasoning have found factors that modify the reasoning process, such as one's utilities (Manktelow & Over, 1991) and probabilities (Kirby, 1994). In this thesis it is taken that deontic reasoning is a form of practical reasoning in which a range of real-world pragmatic effects may be observed.

Traditionally, reasoning has been investigated using logical reasoning tasks that are truth functional. However, more recently there has been a trend to use more realistic materials and the *deduction paradigm* has begun to be questioned (Evans, 2002). The approach in this thesis has been to endorse this recent trend and the materials used in the reported experiments have involved real-world contexts, familiar transgressions, inducements, scaled violations and power relations that would be known to all adult populations.

The distinction between System 1 and System 2 thinking (Stanovich, 1999) is important as it represents everyday thinking and formal, logical thinking, respectively. The cognitive processes for the two forms appear to be quite different but have

traditionally been ignored by theories of reasoning. Pragmatic factors would fit the characteristics of System 1 as they are associated with personal knowledge and experience and can be invoked readily, with little conscious processing. Reasoning pragmatically does not seem to require deduction. However, that does not preclude the possibility of formal, rule-based reasoning occurring.

Research findings with the utilities and probabilities associated with deontic reasoning has lead to the suggestion that reasoning and decision making are invariably linked (Evans, Over & Manktelow, 1993). However, there has been no research to test this hypothesis. The judgement revision experiments reported in this thesis provide support for such a view and suggest there is much scope for developing this strand of research.

The specific pragmatic factors investigated in this thesis have been scale of violation, circumstances, and power of source. Each has been found to modify both reasoning and decision-making processes. These factors are associated with deontic contexts and transgressions of rules and laws. However, there is no reason why aggravating and mitigating circumstances and power of source may not apply to other contexts. Furthermore, the findings here suggest further pragmatic factors should be identifiable.

A number of unpredicted findings were also observed in the series of experiments. The findings of Experiment 1 in which an effect was observed for Day 1 but not Day 2 is difficult to explain. The perception of transgressions also influenced findings; drink-driving was consistently considered more serious than speeding and there were

few circumstances that mitigated against a drink-driving offence. Further, scale of violation varied between the two motoring transgressions with the minor speeding offence barely considered an offence at all, in contrast with drink-driving in which a minor violation was often considered comparable to a major violation. Only by considering the pragmatics of transgressions can these sorts of findings be discovered.

Limitations of the experiments reported in the thesis should also be mentioned. The number and range of social contexts may be considered rather limited, although some consideration was given to the inclusion of formal and informal institutions. However, one advantage of using a reduced range of social contexts is that it allows genuine effects to emerge. For example, a number of studies have used a greater number of scenarios to investigate power (Ray, Raynolds & Carranza, 1989) and inducements (Ohm & Thompson, 2004, 2006). These studies have obtained some contrasting findings to the results reported here and part of the reason may lie in the social contexts of their materials. Ray et al. used some power relations that are questionable and Ohm & Thompson used several inducements that confounded with power of source.

The tasks given to participants varied from study to study, ranging from fines, likelihood ratings of action to credibility ratings and probability estimates. One could argue for a more consistent approach to the measures used to allow for a comparison of effects across experiments. For instance, a better comparison of performance would be possible between reasoning and judgement revision tasks if both measured the same dependent variable. However, the range of measures and tasks does suggest that pragmatic factors produce reliable and valid effects.

The findings with inducements reported in this thesis were not consistent with previous research. All types of inducement were observed to have similar ratings whereas Newstead, Ellis, Evans & Dennis (1997) found subtle differences. Certainly, the definitions of each inducement differed between this thesis and those of Newstead et al. Each inducement (promise, threat, warning etc.) may have several forms requiring further investigation. A similar argument may be applied to conjunctive and disjunctive statements.

The introduction of the superordinate level of statement in Chapter 7 produced comparable findings to those at the non-superordinate level. However, a number of issues arise. It may not always be clear what the superordinate level is. For instance, with the experiments involving drink-driving and speeding, one may state the superordinate level as *motoring offences*. However, clear differences are observed between the two motoring transgressions and the impact of mitigating and aggravating circumstances. Therefore, the theory may require further elaboration to incorporate pragmatic influences.

The approach taken here is that pragmatic factors produce reliable and valid effects in the context of deontic reasoning. Deontic reasoning is considered a type of practical, everyday reasoning involving System 1 thinking which is associated with decision-making and judgement revision. The pragmatics of scale of violation, circumstances and power of source produce modifying influences when presented in familiar, realistic contexts. A number of findings were original and unpredicted and may provide the opportunity for further research.

8.4 Theoretical Discussion of Experimental Findings

This section of the discussion will consider the major theories of reasoning in view of the findings obtained from the experiments with pragmatic factors of deontic reasoning.

The earliest theory of deontic reasoning, Pragmatic Reasoning Schema (PRS) theory (Cheng & Holyoak, 1985), proposed a cognitive schema which is initiated when an obligation or permission is identified. However, the implicit suggestion is that the schema either fires or does not fire; there is no mention in the theory of partial activation as observed with scaled violations or sources of power. A schema-type theory may be appropriate for representing ones knowledge of general permissions and obligations but PRS theory appears inadequate in explaining the pragmatics of reasoning; the theory assumes ones recognition of a permission or obligation and activation of the appropriate schema as a priori. A possible modification of schema theory, to incorporate the effects reported in this thesis, would be to consider each pragmatic factor as an individual schema, however, this would not be a cognitively efficient form of representation.

Evolutionary theories also have difficulty in explaining the subtleties of pragmatic effects on deontic reasoning. Social Contract theory (SCT) may offer a parsimonious explanation of how one reasons with violations and detects violators of social contracts but the theory cannot explain the relative responses to scaled violations or why one violation is considered more serious than another. Two major criticisms can be cast at social contract theory. Firstly, it lacks a semantic (or schematic) component containing acquired knowledge of violations from which to draw relevant knowledge

when reasoning with a range of infringements. Secondly, the activation of the cheater-detection module is presented as an all-or-none event. The result of these deficiencies is that the theory fails to recognise that many laws and rules are arbitrary and invented within a particular social culture to fit a specific demand of the day.

A feature of reasoning with rules is non-monotonicity, that is, drawn conclusions can be modified by subsequent information. The inclusion of pragmatic factors, particularly when presented in series in a judgement revision task, demonstrates reasoning and decision making to be monotonic. This creates particular difficulty for logic-based theories. However, theories based on logic have assumed System 2 thinking while the focus of this thesis has been System 1 thinking.

Ohm & Thompson (2006) in their study of inducements found some overlap between conditional probability and the probability of q, given not-p, and sufficiency and necessity. Perceived sufficiency reflects the extent to which p (the antecedent) guarantees the outcome of q (the consequent) while perceived necessity reflects the extent to which p (the antecedent) is a required condition for q (the consequent) to occur. Sufficient relations would have the $P(q|p)$ close to one and necessary relationships would have the $P(q|\neg p)$ close to zero. The notion of sufficiency and necessity has been developed by a number of authors (Cummins, Lubart, Alksnis & Rist, 1991; Fairley, Manktelow & Over, 1999). In the context of the findings reported in the experiments of this thesis knowing that a motorist is drink-driving is sufficient for them to receive a heavy fine. For a motorist travelling at 35 mph in a 30 mph zone, it would be necessary for additional circumstances to apply for the motorist to receive a fine.

Mental models theory (MMT) (Johnson-Laird & Byrne, 1991) would also have difficulty incorporating the findings reported in this thesis, not least because Johnson-Laird and his many collaborators have concentrated on abstract and arbitrary materials, avoiding the more realistic content as used here. However, a modified version of the theory may be able to explain pragmatic influences, particularly if combined with the probability theory of Evans and colleagues (Evans, Handley & Over, 2003; Evans, Over & Handley, 2003; Evans & Over, 2004).

In this dual theory a statement would be represented by a mental model, incorporating pragmatic knowledge. The probability of the truth or felicity of the statement is then assessed. Mental models theory proposes that only an implicit model of given information is initially devised and true deduction requires the explicit evocation of counter examples. This initial model can be considered as a representation of System 1 thinking; an immediate, relatively unconscious mental representation of the given information. For example, given the statement:

> If a motorist travels above 30mph in a built-up area and is stopped by the police, then he or she is liable to a fine.

The following model may be evoked to represent the statement:

> 30 mph   Fine

If informed that the motorist is travelling above 30 mph then one may draw the modus

ponens conclusion that the motorist is liable to a fine. However, when the speed of the

motorist is specified such a conclusion does not automatically follow. For instance, if

informed the following:

A motorist travels at 35 mph

The following model may be constructed:

35 mph   No Fine

Pragmatic knowledge determines that travelling at 35 mph does not constitute a

violation and is unlikely to result in a fine. Hence, the probability of the statement

being made true is significantly reduced. With this particular case there may be little

need for the inclusion of probabilities and pragmatic knowledge applied to mental

models may be sufficient.

However, cases involving power of source and judgement revision appear more akin

to a probabilistic explanation. For example, given the following inducement:

A boy is promised, "Mow the lawn and I will give you your pocket money".

The inducement may be represented as a mental model with the following:

Mow lawn   Pocket money

If it is then established that the promise has been made by the boy's father, then one is likely to draw the conclusion that the boy will most likely receive his pocket money if he mows the lawn. However, the probability of q (will give you your pocket money) is reduced when it is known that the promise was made by the boys brother. Hence, the probability of the statement being made true, particularly the probability of q|p, is reduced. In mental models terms, the representation may be as follows:

Mow lawn   Pocket money   [Brother]

Mow lawn   No pocket money

Note: [Brother] represents the fact that the inducement has been made by the brother, resulting in a counter example being evoked.

A judgement revision task would involve similar models being devised with revised models formulated as subsequent information is received. Information consistent with the initial statement and model may require no modification, as observed with cases where ratings are the same (e.g. greater violation and high power of source produce high likelihood of action ratings). Contrasting information would require a significant revision of model due to the change in probabilities, as observed with cases where ratings substantially change (e.g. greater violation may produce high likelihood ratings but subsequent information detailing that the statement was made by someone of low authority produces a significant decrease in the likelihood rating).

It should be noted that this proposal advances Manktelow & Over's (1995) suggestion for the inclusion of a '+' and '-' to represent changes to personal preferences or

utilities because more semantic information is retained. The mental models as presented here represent only the information given; implicit models are not made explicit and there is no conscious search for counter examples. As System 1 processing is suggested to be occurring, logical deduction (System 2 thinking) is not required. Pragmatic knowledge can be implicitly inferred within the mental model and probabilistic information can be represented as an additional token. Mental models theory provides a mental representation of the given information which probabilistic accounts do not offer.

The preferred explanation on the grounds of parsimony is the probabilistic theory of Evans et al. (Evans, Handley & Over, 2003; Evans, Over & Handley, 2003; Evans & Over, 2004). Although the theory does not offer a cognitive representation of reasoning and decision-making it does cover a broader scope of the research findings reported here, without the necessity for modification. It is, therefore, proposed that deontic statements and inducements are considered in terms of the conditional probability hypothesis (P(q|p); what is the probability of q being made true, given the antecedent (p). A major violation or high power source are pragmatic influences that are most likely to see q come true. The possibility of counter-examples is considered in terms of the probability of the consequent occurring, given not-p, P(q|¬p) and is more likely to be evoked when a minor offence or a low power source is encountered. The felicity of a statement or inducement is questioned in such cases and the number of counter-examples increases.

This section has discussed a range of possible explanations of the findings and has found schema theories inadequate in their current form. Mental models theory, with

substantial modification, could account for pragmatic factors, however, it would require the need for a dual theory. The preferred explanation is that of the conditional probability hypothesis proposed by Evans and colleagues, which appears to provide a parsimonious theory covering the majority of findings.

8.5 Future Research

Proposals for future research include the possibility of developing the range of pragmatic factors and to extend the established factors to other areas of reasoning and decision-making.

The range of possible pragmatic factors could be increased to include the consequences of a transgression. Just as scale of violation has been observed to influence both reasoning and judgement revision one would expect more serious consequences of a transgression to be rated more highly than less serious consequences. For example, within the context of motoring transgressions a serious collision or accident may produce a greater perception of guilt towards the motorist compared to a minor collision.

To develop the theoretical aspects of the research findings, probability estimates could be taken for a judgement revision task. It would be predicted that revised judgements would be aligned to the probabilities observed to date; major offence and high power source would increase the probability of action while a minor offence and low power source would lower probability estimates. This would also provide a suitable means of testing the application of the conditional probability hypothesis to the area of judgement revision.

Manipulations of the current findings could also be undertaken. The perspective effect is well established with deontic reasoning tasks and could be applied to the felicity of statements and inducements. For instance, a Sergeant (typically high power) may have authority when advising a teenager about a career in the army but have little authority when advising him about an illness; conversely, a teenager may take the advice of a doctor about illnesses but not about careers in the army. There is the possibility that such manipulations would be observed with a thematic version of the selection task.

The experiments reported in this thesis and the identified pragmatic factors have been confined to deontic reasoning contexts. However, there may be the possibility for the pragmatics to be applied to other types of thinking, such as causal reasoning. For instance, given there has been a car accident and knowing that a driver has been drinking may be sufficient to imply that they are guilty of the accident. Indeed, it is common practice for police officers to breathalyse drivers at the scene of an accident in order to aid the establishment of the causes of the event.

The research undertaken for this thesis provides opportunities to further expand the exciting work of pragmatics of deontic reasoning and to develop observed pragmatic effects to other areas of reasoning, such as causal reasoning.

References

Almor, A., & Sloman, S.A. (1996). Is deontic reasoning special? *Psychological Rview, 103,* 374-380.

Anderson, J.R. (1990). *The adaptive character of thought*. Hillside, NJ: Lawrence Erlbaum.

Anderson, N.H. (1970). Functional measurement and psychophysical judgment. *Psychological Review, 77,* 153-170.

Anderson, N.H. (1974). Information integration theory: A brief survey. In D.H. Krantz, R.C. Atkinson, R.D. Luce & P. Suppes (Eds), Contemporary developments in mathematical psychology (Vol.2). San Francisco: Freeman.

Anderson, N.H. (1981). Foundations of integration theory. New York: Academic Press.

Asch, S.E. (1946) Forming impressions of personality. *Journal of Abnormal & Social Psychology, 4,* 258-290.

Audi, R. (1989). *Practical Reasoning*. London: Routledge & Kegan Paul.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Axelrod, R., & Hamilton, W.D. (1981). The evolution of cooperation. *Science, 211,* 1390-1396.

Beach, L.R. (1997). *The psychology of decision making: people in organizations*. London: Sage.

Behn, R.D., & Vaupel, J.W. (1982). *Quick analysis for busy decision makers.* New York: Basic Books.

Bernoulli, D. (1738/ 1954). Specimen theoriae novae de mensura sortis. *Comentarii Academiae Scieniarum Imperiales Petropolitanae, 5,* 175-192.

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.

Bochner, S., & Insko, C.A. (1966). Communicator discrepancy, source credibility, and opinion change. *Journal of Personality and Social Psychology*, 4, 614-621.

Bonatti, L. (1994). Propositional reasoning by model? *Psycholgical Review, 101*, 725-733.

Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85,* 1-21.

Braine, M.D.S., Reiser, B.J., & Rumain, B. (1984). Some empirical justification fpr a

theory of mental propositional logic. In G. Bower (Ed), *The psychology of learning and motivation: Advances in research and theory* (Vol 18). New York: Academic Press.

Braine, M.D.S., & Rumain, B. (1981). Development of comprehension of "or": Evidence for a series of competencies. *Journal of Experimental Child Psychology, 31,* 46-70.

Brehmer, B. (1990). Strategies in real-time, dynamic decision making. In R. Hogarth (Ed), *Insights in decision making: A tribute to the late Hilel J. Einhorn.* Chicago, IL: University of Chicago Press.

Brehmer, B., & Brehmer, A. (1988). What have we learned about human judgement from thirty years of policy capturing? In B. Brehmer & C.R.B. Joyce (Eds), *Human judgement: The SJT view.* Amsterdam: North-Holland.

Brown, R.V., Kahr, A.S., & Peterson, C.R. (1974). *Decision analysis for the manager.* New York: Holt, Rinehart & Winston.

Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking.* New York: Wiley.

Brunswick, E. (1937). Psychology as a science of objective relations. *Philosophy of Science, 4,* 227-260.

Brunswick, E. (1947). *Systematic and representative design of psychological experiments, with results in physical and social perception.* Berkeley: University of California Press.

Brunswick, E. (1952). The conceptual framework of psychology. In *International encyclopedia of unified science* (Vol. 1, No. 10). Chicago, IL: University of Chicago Press.

Byrne, R.M.J. (1989). Supressing valid inferences with conditionals. *Cognition, 31,* 61-83.

Byrne, R.M.J., & Johnson-Laird, P.N. (1989). Spatial reasoning. *Journal of Memory and Language, 28,* 564-575.

Camerer, C.F. (1981). General conditions for the success of bootstrapping models. *Organizational Behavior and Human Performance, 27,* 411-422.

Chapman, L.J. & Chapman, J.P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology. 58,* 220-226.

Chater, N., & Oaksford, M. (1996). Deontic reasoning, modules and innateness: A second look. *Mind and Language, 11,* 191-202.

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38(2),* 191-258.

Chellas, B.F. (1980). Modal logic: an introduction. Cambridge, UK: Cambridge University Press.

Cheng, P.W., & Holyoak, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391-416.

Cheng, P.W., & Holyoak, K.J., Nisbett, R.E., & Oliver, L.M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293-328.

Cheng, P.W., & Novick, L.R. (1991). Causes versus enabling conditions. *Cognition, 40,* 83-120.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-316.

Cosmides, L. & Tooby, J. (1989). Evolutionary psychology and the generation of culture, Part II. Case study: A computational theory of social exchange. *Ethology and Sociobiology, 10,* 51-97.

Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds), *The adapted mind.* Oxford, UK: Oxford University Press.

Cosmides, L. & Tooby (1994). Beyond intuition and instinct blindness: toward an evolutionary rigorous cognitive science. *Cognition, 50,* 41-77.

Craik, K. (1943). *The nature of explanation.* Cambridge, UK: Cambridge University Press.

Cummins, D.D. (1995). Naïve theories and causal deduction. *Memory and Cognition, 23,* 646-658.

Cummins, D.D. (1996a). Evidence for the innateness of deontic reasoning. *Mind & Language, 11,* 160-190.

Cummins, D.D. (1996b). Evidence of deontic reasoning in 3- & 4-year-old children. *Memory & Cognition, 24,* 823-829.

Cummins, D.D. (1998). Social norms and other minds: The evolutionary roots of higher cognition. In D.D. Cummins & C. Allen (Eds), *The evolution of mind.* Oxford, UK: Oxford University Press.

Cummins, D.D. (1999). Cheater detection is modified by social rank: The impact of dominance on the evolution of cognitive functions. *Evolution and Human Behavior,* 20, 229-248.

Cummins, D.D. (2000). How the social environment shaped the evolution of mind. *Synthese, 122,* 3-28.

Cummins, D.D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition, 19,* 274-282.

Dawes, R.M. (1971). A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist, 26,* 180-188.

Dawes, R.M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34,* 571-582.

Dawes, R.M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81,* 95-106.

de Finetti, B. (1937). La prevision: Ses lois logiques, ses sources subjectives [Foresight: Its logical laws, its subjective sources]. *Annales de l'InstituteHenri Poincare, 7,* 1-68. Translated in H.E. Kyburg & H.E. Smokler (Eds) (1964), *Studies in subjective probability.* Chichester: Wiley.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2003). Inference suppression and semantic memory retrieval: every counterexample counts. *Memory and Cognition, 31*, 581-595.

De Paulo, B.M., & Coleman, L.M. (1986). Talking to children, foreigners and retarded adults. *Journal of Personality and Social Psychology, 51,* 945-959.

De Paulo, B.M., & Coleman, L.M. (1987). Verbal and nonverbal communication of warmth to children, foreigners, and retarded adults. *Journal of Nonverbal Behaviour, 11,* 75-88.

Doherty, M.E., & Kurz, E.M. (1996). Social judgement theory. *Thinking and Reasoning, 2 (2/3),* 109-140.

Dominowski, R.L. (1995). Content effects in Wason's selection task. In S.E. Newstead & J.St.B.T. Evans (Eds), *Perspectives on thinking and reasoning. Essays in honour of Peter Wason.* Hove: Lawrence Erlbaum Associates Ltd.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51,* 380-417.

Edwards, W., & Newman, J.R. (1982). *Multiattribute evaluation.* Beverly Hills: Sage.

Einhorn, H. & Hogarth, R. (1975). Unit weighting schemes for decision making. *Organizational Behaviour and Human Performance, 13*, 171-192.

Einhorn, H. & Hogarth, R. (1981). Behavioral decision theory: processes of judgement and choice. *Annual Review of Psychology, 31,* 53-88.

Erber, R. & Fiske, S.T. (1984). Outcome dependency and attention to inconsistent information about others. *Journal of Personality and Social Psychology, 47,* 709-726.

Evans, J.St.B.T. (1982). *The psychology of deductive reasoning.* London: Routledge

& Kegan Paul.

Evans, J.St.B.T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75,* 451-468.

Evans, J.St.B.T. (1989). *Bias in human reasoning: causes and consequences.* Hove: Lawrence Erlbaum Associates.

Evans, J.St.B.T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin,* 128, 978-996.

Evans, J.St.B.T., Barston, J.L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition, 11,* 295-306.

Evans, J.St.B.T., Handley, S.J., Hadjichristidis, C., Thompson, V., Over, D.E., & Bennett, S. (2007). On the basis of belief in causal and diagnostic conditionals. *Quarterly Journal of Experimental Psychology, 60,* 635-643.

Evans, J.St.B.T., Handley, S.J., & Over, D.E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 321-335.

Evans, J.St.B.T., Legrenzi, P., & Girotto, V. (1999). The influence of linguistic form on reasoning: the case of matching bias. Quarterly *Journal of Experimental Psychology, 52A,* 185-216.

Evans, J.St.B.T., & Lynch, J.S. (1973). Matching bias in the selection task. *British Journal of Psychology, 64,* 391-397.

Evans, J.St.B.T., Newstead, S.E. & Byrne, R.M.J. (1993). *Human reasoning: The psychology of deduction.* Hove: Lawrence Erlbaum Associates Ltd.

Evans, J.St.B.T., & Over, D.E. (1996a). Rationality in the selection task: epistemic utility versus uncertainty reduction. *Psychological Review, 103,* 356-363.

Evans, J.St.B.T., & Over, D.E. (1996b). *Rationality and reasoning.* Hove: Erlbaum.

Evans, J.St.B.T., & Over, D.E., & Handley, S.J. (2003). A theory of hypothetical thinking. In D. Hardman & L. Macchi (Eds), *Thinking: psychological perspectives on reasoning, judgment and decision making.* Sussex: J. Wiley & Sons Ltd.

Evans, J.St.B.T., & Over, D.E. (2004). *If.* Oxford: Oxford University Press.

Evans, J.St.B.T., Over, D.E., & Handley, S.J. (2005). Suppositions, extensionality, and conditionals: A critique of the mental model theory of Johnson-Laird and Byrne (2002). *Psychological Review, 112,* 1040-1052.

Evans, J.St.B.T., Over, D.E., & Manktelow, K.I. (1993). Reasoning, decision making and rationality. *Cognition, 49,* 165-187.

Evans, J.St.B.T., & Twyman-Musgrove, J. (1998). Conditional reasoning with inducements and advice. *Cognition, 69,* B11-B16.

Fairley, N., & Manktelow, K.I. (1997). Causal and conditional reasoning: A comment on Cummins (1995). *Memory and Cognition, 25,* 413-414.

Fairley, N., Manktelow, K.I., & Over, D.E. (1999). Necessity, sufficiency, and perspective effects in causal and conditional reasoning. *Quarterly Journal of Experimental Psychology, 52A,* 771-790.

Fairley, N., & Manktelow, K.I. (2004). Superordinate principles, conditions and conditionals. In D. Hardman & L. Macchi (Eds), *Thinking: psychological perspectives on reasoning, judgment and decision making.* Sussex: J. Wiley & Sons Ltd.

Feeney, A., Evans, J.St.B.T., & Clibbens, J. (1996). Information order, expectation & pragmatics: the case of sequential pseudodiagnostic evidence. *Poster presentation at The Third International Conference on Thinking, UCL, August.*

Fillenbaum, S. (1975). If: Some uses. *Psychological Research, 37,* 245-260.

Fillenbaum, S. (1976). Inducements: on phrasing and logic of conditional promises, threats and warnings. *Psychological Research, 38,* 231-250.

French, J.R.P., & Raven, B.H. (1959). The bases of social power. In D. Cartwright (Ed), *Studies in social power.* Ann Arbor, MI: Institute for Social Research.

Garcia-Madruga, J.A., Moreno, S., Carriedo, N., Gutierrez, F. & Johnson-Laird, P.N. (2001). Are conjunctive inferences easier than disjunctive inferences? A comparison of rules and models. *Quarterly Journal of Experimental Psychology, 54A (2),* 613-632.

George, C. (1995). The endorsement of the premises: Assumption-based or belief-based reasoning. *British Journal of Psychology*, 86, 93-111.

George, C. (1997). Reasoning from uncertain premises. *Thinking and Reasoning*, 3, 161-189.

Gigerenzer, G., & Goldstein, D.G. (1996). Reaonin gthe fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650-669.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition, 43,* 127-171.

Gigerenzer, G., Todd, P.M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilhooly, K.J., & Falconer, W.A. (1974). Concrete and abstract terms and relations in testing a rule. *Quarterly Journal of Experimental Psychology, 26,* 355-359.

Girotto, V. (1995). Contextual factors ob deontic reasoning. *Thinking and Reasoning,*

1, 333-339.

Girotto, V., Blaye, A., & Farioli, F. (1989). A reason to reason: Pragmatic bases of children's search for counter examples. *European Bulletin of Cognitive Psychology, 9,* 227-231.

Girotto, V., Light, P., & Colbourn, C.J. (1988). Pragmatic schemas and conditional reasoning in children. *Quarterly Journal of Experimental Psychology, 40A,* 469-482.

Girotto, V., Mazzocco, A., & Tasso, A. (1997). The effect of premise order in conditional reasoning: A test of the   mental model theory. *Cognition, 63,* 1-28.

Goldberg, L.R. (1970). Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inference. *Psychological Bulletin, 73,* 422-432.

Green, D.W., McClelland, A., Muckli, L., & Simmons, C. (1999) Arguments and deontic decision. *Acta Psychologica, 101*, 27-47.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan. *Syntax and semantics: Speech acts.* New York: Academic Press.

Griggs, R.A., & Cox, J.A. (1982). The elusive thematic materials effect in the Wason selection task. *British Journal of Psychology, 73,* 407-420.

Hacker, H.M. (1951). Women as a minority group. *Social Forces, 30,* 60-69.

Hammond, K.R. (1955). Probabilistic functioning and the clinical method. *Psychological Review, 62,* 255-262.

Hammond, K.R. (1986). *A theoretically based review of theory and research in judgment and decision making.* (Technical report No. 260). Boulder: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K.R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice.* New York: Oxford University Press.

Hastie, R. (2001). Problems for judgement and decision making. *Annual Review of Psychology, 52,* 653-683.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review, 69,* 366-378.

Hogarth, R.M., & Einhorn, H.J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24,* 1-55.

Hogarth, R.M., & Karelaia, N. (2007). Heuristic and linear models of judgement: Matching rules and environments. *Psychological Review, 114*, 733-758.

Holyoak, K.J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13,* 295-355.

Hurford, J.R. (1974). Exclusive or inclusive disjunction. *Foundations of Language, 11,* 409-411.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking*. New York: Basic Books.

James, W. (1890). The principles of psychology. New York: Holt.

Janis, I., & Frick, P. (1943). The relationship between attitudes toward conclusions and errors in judging logical validity of syllogisms. *Journal of Experimental Psychology, 33,* 73-77.

Johnson-Laird, P.N. (1975). Models of deduction. In R. Falmagne (Ed), *Reasoning: Representation and process.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Johnson-Laird, P.N., & Byrne, R.M.J. (2002). Conditionals: A theory of meaning, pragmatics and inference. *Psychological Review, 109*, 646-678.

Johnson-Laird, P.N., Byrne, R.M.J. & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99(3),* 418-439.

Johnson-Laird, P.N., Legrenzi, P., Girotto, V., Legrenzi, M.S., & Caverni, J.-P. (1999). Naïve probability: A model theory of extensional reasoning. *Psychological review, 106,* 62-88.

Johnson-Laird, P.N., Legrenzi, P., & Legrenzi, M.S. (1972). Reasoning and a sense of reality. *British Journal of Psychology, 63,* 395-400.

Johnson-Laird, P.N., Savary, F., & Bucciarelli, M. (2000). Strategies and tactics in reasoning. In G. De Vooght, G. D'Ydewalle, W. Schaeken, & A. Vandierendonck (Eds.), *Deductive reasoning and strategies*. Mahwah, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P.N., & Wason, P.C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology,1,* 134-148.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgement of representativeness. *Cognitive Psychology, 3,* 430-454.

Kahneman, D., Slovic, P., & Tversky, A. (1982) *Judgement under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Kant, I. (1785/1991). *The metaphysics of morals*. Cambridge: Cambridge University Press.

Keane, M. (1990). Incremental analogising: theory and model. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie, & G. Erdos (eds), *Lines of thought: Reflections on the psychology of thinking, Vol. 1.* Chichester, UK: Wiley.

Keltner, D., Gruenfeld, D.H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110,* 265-284.

Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value trade-offs.* New York: Wiley.

Kirby, K.N. (1994). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition, 51,* 1-28.

Klayman, J. (1988). On the how and why (not) of learning from outcomes. In B. Brehmer & C.R.B. Joyce (Eds), *Human judgement: The SJT view.* Amsterdam: North-Holland.

Kort, F. (1968). A nonlinear model for the analysis of judicial decisions. *American Political Science Review, 62,* 546-555.

Krantz, D.H., Luce, R.D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). New York: Academic Press.

Lakoff, R. (1971). If's and's and but's about conjunction. In C.J. Fillmore & D.T. Langedoen (Eds), *Studies in linguistic semantics.* New York: Holt, Rinehart and Winston.

Laming, D. (1996). On the analysis of irrational data selection: A critique of Oaksford & Chater (1994). *Psychological review, 103(2),* 364-373.

Lefford, A. (1946). The influence of emotional subject matter on logical reasoning. *Journal of General Psychology, 34,* 127-151.

Legrenzi, P., Girotto, V., & Johnson-Laird, P.N. (1993). Focusing in reasoning and decision making. *Cognition, 49,* 37-66.

Levinson, S. (1983). *Pragmatics.* Cambridge: Cambridge University Press.

Lopes, L.L. (1996). When time is of the essence: Averaging, aspiration, and the short run. *Organizational Behavior and Human Decision Processes, 65,* 179-189.

Luchins, A.S. (1957). Primacy-recency in impression formation. In C. Hovland (Ed.), *The order of presentation in persuasion.* New Haven, Conneticut: Yale University Press.

Manktelow, K.I. (1999). *Reasoning and thinking.* Hove: Psychology Press.

Manktelow, K.I. (2004). Reasoning and rationality: The pure and the practical. In K.I. Manktelow & M.C. Chung (Eds.) *Psychology of reasoning: Theoretical and historical perspectives.* Hove: Psychology Press.

Manktelow, K.I. & Evans, J.St.B.T. (1979). Facilitation of reasoning by realism: Effect or non-effect. *British Journal of Psychology, 70,* 477-488.

Manktelow, K.I. & Fairley, N. (2000). Superordinate principles in reasoning with causal and deontic conditionals. *Thinking and Reasoning, 6,* 41-65.

Manktelow, K.I. & Over, D.E. (1990). Deontic thought and the selection task. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie & G. Erdos (eds.), *Lines of thought: Reflections on the psychology of thinking, Vol. 1.* Chichester: Wiley.

Manktelow, K.I. & Over, D.E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition, 39,* 85-105.

Manktelow, K.I. & Over, D.E. (1995). Deontic reasoning. In S.E. Newstead & J.St.B.T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason. Hove,* UK: LEA Ltd.

Manktelow, K.I., Sutherland, E.J. & Over, D.E. (1995). Probabilistic factors in deontic reasoning. *Thinking & Reasoning, 1(3),* 201-219.

Marr, D. (1982). *Vision.* San Francisco: W.H. Freeman.

McClelland, J., Rumelhart, D.E., & the PDP Research Group (1986). *Parallel distributed processing: explorations in the microstructure of cognition. Vol. 2: Psychological & biological models.* Cambridge: MIT Press.

McGuire, W.J. (1960). A syllogistic analysis of cognitive relationships. In M.J. Rosenberg et al. (Eds), *Attitude organization and change.* Yale: Yale University Press.

Mey, J.L. (1993). Pragmatics: An introduction. Oxford: Blackwell.

Mill, J.S. (1861). *Utilitarianism.* London: Collins.

Morgan, J.I.B., & Morton, J.T. (1944). The distortions of syllogistic reasoning produced by personal connections. *Journal of Social Psychology, 20,* 39-59.

Morris, C. (1938). Foundations of the theory of signs. In O. Neurath, R. Carnap, & C. Morris (Eds.), *International encyclopedia of unified science,* 77-138. Chigago: University of Chicago Press. Reprinted as Morris, C. (1971). *Writings in the general theory of signs.* The Hague: Mouton.

Neuberg, S.L., & Fiske, S.T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology, 53,* 431-444.

Neumann, J. von, & Morgenstern, O. (1944). *Theory of games and economic behaviour.* Princeton, NJ: Princeton University Press.

Neumann, J. von, & Morgenstern, O. (1947). *Theory of games and economic*

behaviour (2^(nd) ed.). Princeton, NJ: Princeton University Press.

Newell, A., & Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice –Hall.

Newstead, S.E. (1989). Interpretational errors in syllogistic reasoning. *Journal of Memory and Language, 28*, 78-91.

Newstead, S.E. (1990). Conversion in syllogistic reasoning. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie & G. Erdos (eds.), *Lines of thought: Reflections on the psychology of thinking, Vol. 1.* Chichester: Wiley.

Newstead, S.E. (1995). Gricean implicatures and syllogistic reasoning. *Journal of Memory and Language, 34*, 644-664.

Newstead, S.E., Ellis, M.C., Evans, J.St.B.T., & Dennis, I. (1997). Conditional reasoning with realistic material. *Thinking and Reasoning, 3(1),* 49-76.

Newstead, S.E., Griggs, R.A., & Chrostowski, J.J. (1984). Reasoning with realistic disjunctives. *Quarterly Journal of Experimental Psychology, 36A,* 611-627.

Ng, S.H., Bell, D., & Brook, M. (1993). Gaining turns and achieving high influence in small conversational groups. *British Journal of Social Psychology, 32,* 265-275.

Oaksford, M. & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101(4),* 608-631.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 883-889.

Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: conditional probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 680-693.

O'Brien, D.P. (1995). On the relationship between pragmatic schemas and mental logic. Thinking and Reasoning, 1, 357-364.

O'Brien, D.P., Braine, M.D.S., & Yang, Y. (1994). Propositional reasoning by mental models? Simple to refute in principle and in practice. *Psychological Review, 101*, 711-724.

Ohm, E., & Thompson, V.A. (2004). Everyday reasoning with inducements and advice. *Thinking and Reasoning*, 10, 241-272.

Ohm, E., & Thompson, V.A. (2006). Conditional probability and pragmatic conditionals: Dissociating truth effectiveness. *Thinking and Reasoning*, 12, 257-280.

Over, D.E., Hadjichristidis, C., Evans, J.St.B.T., Handley, S.J., & Sloman, S.A.

(2007). The probability of causal conditionals. *Cognitive Psychology*, 24, 62-97.

Over, D.E., Manktelow, K.I., & Hadjichristidis, C. (2004). Conditions for the acceptance of deontic conditionals. *Canadian Journal of Experimental Psychology, 58,* 96-105.

Pascal, B. (1670/1941). *Pensees.* [Translated by W.F. Trotter]. New York: Modern Library.

Pelletier, F.J. (1977). Or. *Theoretical Linguistics, 4,* 61-74.

Phelps, R.H., & Shanteau, J. (1978). Livestock judges: How much information can an expert use? *Organizational behavior and Human Performance, 21,* 209-219.

Pliskin, J.S., Shepard, D.S., & Weinstein, M.C. (1980). Utility functions for life years and health status. *Operations Research, 28,* 206-224.

Politzer, G., & Nguyen-Xuan, A. (1992). Reasoning about promises and warnings: Darwinian algorithms, mental models, relevance judgements or pragmatic schemas? *Quarterly Journal of Experimental Psychology, 44A,* 402-421.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3,* 323-343.

Ray, J.L., Reynolds, R.A., & Carranza, E. (1989). Understanding choice utterances. *The Quarterly Journal of Experimental Psychology, 41A (4),* 829-848.

Ramsey, F.P. (1931). *The foundations of mathematics and other logical essays.* London: Routledge & Kegan Paul.

Raven, B.H. (1965). Social influence and power. In D. Steiner & M. Fishbein (Eds), *Current studies in social psychology.* New York: Holt, Reinhart & Winston.

Reich, S.S. & Ruth, P. (1982). Wason's selection task: Verification, falsification, and matching. *British Journal of Psychology, 73,* 395-405.

Revlin, R., Leirer, V., Yopp, H., & Yopp, R. (1980). The belief bias effect in formal reasoning: The influence of knowledge on logic. *Memory & Cognition, 8,* 584-592.

Rips, L.J. (1983). Cognitive processes in propositional reasoning. *Psychological Review, 90,* 38-71.

Rodin, M.J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition, 5,* 144-165.

Roose, J.E., & Doherty, M.E. (1976). Judgement theory applied to the selection of life insurance salesman. *Organizational Behavior and Human Performance, 16,* 231-249.

Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension.*

Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Ruscher, J.B., & Fiske, S.T. (1990). Interpersonal competition can cause individuating impression formation. *Journal of Personality and Social Psychology, 58,* 832-842.

Rusk, M.C., & Russell, J.E.A. (1988). Leader prototypes and prototype-contingent consensus in leader behaviour descriptions. *Journal of Experimental Social Psychology, 24,* 88-104.

Sande, G.N., Ellard, J.H., & Ross, M. (1986). Effect of arbitrarily assigned status labels on self-perceptions and social perceptions: The mere position effect. *Journal of Personality and Social Psychology, 50,* 684-689.

Savage, L. (1954). *The foundations of statistics*. New York: Wiley.

Searle, J.R. (1969). *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.

Shafir, E.B. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory and Cognition, 21,* 546-556.

Shafir, E.B., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology, 24,* 449-474.

Simon, H.A. (1982). *Models of bounded rationality.* Cambridge, MA: MIT Press.

Simon, H.A. (1983). *Reason in human affairs.* Stanford: Stanford University Press.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119(1),* 3-22.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology, 28,* 1-39.

Slovic, P., Rorer, L.G., & Hoffman, P.J. (1971). Analyzing the use of diagnostic drugs. *Investigative Radiology, 6,* 18-26.

Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57,* 31-95.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Oxford: Blackwell.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition (2nd ed.).* Oxford: Blackwell.

Springston, F.J., & Clark, H.H. (1973). "And" and "or", or the comprehension of pseudoimperatives. *Journal of Verbal Learning and Verbal Behaviour, 12,* 258-272.

Stanovich, K.E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Stanovich, K.E., & West, R.F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89,* 342-357.

Stanovich, K.E., & West, R.F. (1998). Cognitive ability and variation in selection task performance. *Thinking and Reasoning, 4,* 193-230.

Stanovich, K.E., & West, R.F. (2000). Individual differences in reasoning: implications for the rationality debate. *Behavioral and Brain Sciences, 23,* 645-726.

Stenning, K., & Oaksford, M.R. (1993). Rational reasoning and human implementations of logic. In K.I. Manktelow & D.E. Over (Eds.), *Rationality: Psychological and Philosophical perspectives*. London: Routledge.

Stevenson, R.J., & Over, D.E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology, 48A(3),* 613-643.

Stevenson, R.J., & Over, D.E. (2001). Reasoning from uncertain premises: Effects of expertise and controversial content. *Thinking and Reasoning, 7,* 367-390.

Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46,* 35-57.

Tversky, (1969). Intransitivity of preferences. *Psychological Review, 76,* 31-48.

Tversky A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Tversky A., & Kahneman, D. (1974). Judgement under uncertainty: heuristics and biases. *Science,185,* 1124-1131.

Tversky A., & Shafir, E.B. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 3,* 305-309.

van der Henst, J-B. (1999). The mental model theory of spatial reasoning re-examined: The role of relevance in premise order. *British Journal of Psychology, 90,* 73-84.

Van Duyne, P.C. (1976). Necessity and contingency in reasoning. *Acta Psychologica, 40,* 85-101.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research.* Cambridge: Cambridge University Press.

Wason, P.C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129-140.

Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20,* 273-281.

Wason, P.C., & Brooks, P.G. (1979). THOG: The anatomy of a problem. *Psychological Research, 41,* 79-90.

Wason, P.C., & Green, D. (1984). Reasoning and mental representation. *Quarterly Journal of Experimental Psychology, 36A,*597-610.

Wason, P.C., & Johnson-Laird, P.N. (1972). *Psychology of reasoning: Structure and content.* London: Batsford.

Wason, P.C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology, 23,* 63-71.

Wigton, R.S. (1996). Social judgement theory and medical judgement. *Thinking and Reasoning, 2(2/3),* 175-190.

Wilkins, M.C. (1928). The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology,* No. 102.

Woodworth, R.S., & Sells, S.B. (1935). An atmosphere effect in syllogistic reasoning. *Journal of Experimental Psychology, 18,* 451-460.